# Psychophysiological evidence of a role for emotion in the learning of trustworthiness from identity-contingent eye-gaze cues.

Luis Richard Manssuer

This thesis is submitted to the School of Psychology, Bangor University, Wales, in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

# Declaration and Consent

## Details of the Work

I hereby agree to deposit the following item in the digital repository maintained by Bangor University and/or in any other repository authorized for use by Bangor University.

**Author Name:** Luis Manssuer

**Title:** Psychphysiological evidence of a role for emotion in the learning of trustworthiness from identity-contingent eye-gaze cues.

**Supervisor/Department:** Professor Steven Tipper, Department of Psychology, University of York.

**Funding body (if any):** School of Psychology, Bangor University

**Qualification/Degree obtained:** Doctor of Philosophy

This item is a product of my own research endeavours and is covered by the agreement below in which the item is referred to as "the Work". It is identical in content to that deposited in the Library, subject to point 4 below.

## Non-exclusive Rights

Rights granted to the digital repository through this agreement are entirely non-exclusive. I am free to publish the Work in its present version or future versions elsewhere.

I agree that Bangor University may electronically store, copy or translate the Work to any approved medium or format for the purpose of future preservation and accessibility. Bangor University is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

## Bangor University Digital Repository

I understand that work deposited in the digital repository will be accessible to a wide variety of people and institutions, including automated agents and search engines via the World Wide Web.

I understand that once the Work is deposited, the item and its metadata may be incorporated into public access catalogues or services, national databases of electronic theses and dissertations such as the British Library's EThOS or any service provided by the National Library of Wales.

I understand that the Work may be made available via the National Library of Wales Online Electronic Theses Service under the declared terms and conditions of use (http://www.llgc.org.uk/index.php?id=4676). I agree that as part of this service the National Library of Wales may electronically store, copy or convert the Work to any approved medium or format for the purpose of future preservation and accessibility. The National Library of Wales is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

**Statement 1:**

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless as agreed by the University for approved dual awards.

Signed ………………………………………….. (candidate)

Date ……………………………………………..

**Statement 2:**

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

All other sources are acknowledged by footnotes and/or a bibliography.

Signed …………………………………………. (candidate)

Date …………………………………………….

**Statement 3:**

I hereby give consent for my thesis, if accepted, to be available for photocopying, for inter-library loan and for electronic repositories, and for the title and summary to be made available to outside organisations.

Signed …………………………………………. (candidate)

Date …………………………………………….

**Statement 4:**

Choose **one** of the following options

| | | |
|---|---|---|
| a) | I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University and where necessary have gained the required permissions for the use of third party material. | Yes |
| b) | I agree to deposit an electronic copy of my thesis (the Work) in the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University when the approved **bar on access** has been lifted. | |
| c) | I agree to submit my thesis (the Work) electronically via Bangor University's e-submission system, however I **opt-out** of the electronic deposit to the Bangor University (BU) Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorized for use by Bangor University, due to lack of permissions for use of third party material. | |

*Options B should only be used if a bar on access has been approved by the University.*

**In addition to the above I also agree to the following:**

1. That I am the author or have the authority of the author(s) to make this agreement and do hereby give Bangor University the right to make available the Work in the way described above.
2. That the electronic copy of the Work deposited in the digital repository and covered by this agreement, is identical in content to the paper copy of the Work deposited in the Bangor University Library, subject to point 4 below.
3. That I have exercised reasonable care to ensure that the Work is original and, to the best of my knowledge, does not breach any laws – including those relating to defamation, libel and copyright.
4. That I have, in instances where the intellectual property of other authors or copyright holders is included in the Work, and where appropriate, gained explicit permission for the inclusion of that material in the Work, and in the electronic form of the Work as accessed through the open access digital repository, *or* that I have identified and removed that material for which adequate and appropriate permission has not been obtained and which will be inaccessible via the digital repository.
5. That Bangor University does not hold any obligation to take legal action on behalf of the Depositor, or other rights holders, in the event of a breach of intellectual property rights, or any other right, in the material deposited.
6. That I will indemnify and keep indemnified Bangor University and the National Library of Wales from and against any loss, liability, claim or damage, including without limitation any related legal fees and court costs (on a full indemnity bases), related to any breach by myself of any term of this agreement.

Signature: …………………………………………………… Date : ……………………………….

# Acknowledgements

I am eternally grateful to Professor Steve Tipper for giving me the opportunity to pursue this PhD and for his unwavering support throughout. I would also like to thank anyone else at Bangor and beyond who has helped in any way along the journey but particularly Ralph Pawling, Amy Hayes and Mark Roberts for their advice and technical assistance.

## Contents

# Chapter 1

## Thesis Summary

There is an increasing recognition that emotion influences cognition. This is particularly clear in the domain of social cognition where the perception of social cues is most often a potent source of emotional arousal. One process likely to be influenced by emotion is the forming of face evaluations. Judgements of trustworthiness are particularly important due to the potential costs and benefits of the decision to rely upon another person. Face trustworthiness is modulated by one particular social cue, eye-gaze. In reaction time tasks, when faces gaze away from target objects, incongruently, responses are slower than when faces gaze towards target objects, congruently. Faces that consistently gaze incongruently are also judged less trustworthy than faces that consistently gaze congruently.

In six experiments, we investigated the role of emotion in the learning of trustworthiness from identity-contingent gaze-cues using event-related potentials (ERP) and facial electromyography (EMG), which are sensitive electrophysiological measures of emotion. We found that the learning of trust was paralleled by an increase in the emotion-related late positive potential (LPP) to incongruent faces across blocks. These findings were further supported by EMG measurements, which showed that corrugator muscle activity related to negative embodied emotional states, was greater to incongruent faces in those participants who showed expected changes in trust ratings. Although effects of gaze-cues on trust were consistent, they were not modulated by extremes in initial trust or priming of emotions elicited by social exclusion. Effects appeared to be due to a special relationship between faces, gaze, emotion and trust as the effects of validity on liking ratings were much weaker or non-existent compared to trust despite evidence of similar emotion-related LPP ERP and EMG activity. The effects also did not generalise to non-social arrow cues. In sum, we conclude that emotion mediates the learning of trust from identity-contingent gaze-cues.

## Introduction

### Origins and Essential Background

The origins of the research, which form the basis for further exploration in this thesis, begin with early evidence for covert spatial cueing of attention. Posner (1980; Posner & Cohen, 1984) showed that attention could be cued endogenously (voluntarily) by a centrally presented arrow stimulus or exogenously (automatically, reflexively) by a sudden onset stimulus presented in the periphery. Reaction times (RTs) to targets presented laterally to participants' centre of fixation were faster when either a central arrow pointed towards the target or when a peripheral cue briefly appeared at the location of the target (congruent trials) compared to when no cue was presented (neutral trials) or when the cue was in the opposite direction to the targets location (incongruent trials). Such effects typically peak when the stimulus onset asynchrony (SOA) between the cue and the target is set at around 100 ms for peripheral cues and at around 300 ms for central cues (Cheal & Lyon, 1991). However, for peripheral cues, when the SOA is much longer, say 500 ms, the cueing effect becomes reversed and participants are faster to respond on incongruent trials compared to congruent trials (Posner & Cohen, 1984). This effect is known as inhibition of return (IOR) and highlights two fundamental exogenous attentional processes. Firstly, attention is directed toward the location of the cue. If the target appears at that location after a short amount of time (~100 ms), RTs are facilitated compared to when incongruent. Secondly, if no target is presented at the cued location within a short amount of time (~500 ms SOA), attention at that location is then inhibited to facilitate detection of stimuli in other locations. As such, RTs become slower on congruent trials due to the need to reorient attention back to the inhibited location. In this way IOR makes the attentional system more sensitive to new stimuli. Although early experiments suggested differences between central arrow cues and peripheral cues in the ability to cue attention exogenously, subsequent studies have shown that central arrow cues can also cue attention exogenously (Tipples, 2002; 2008).

This basic attention research paradigm provided an ideal means with which to investigate a long standing recognition for the role of gaze perception and gaze following in social cognition and development (Baron-Cohen, 1995). In particular, Baron-Cohen (1995) proposed that gaze following and the joint attention that results is automatic, emerges early in development (14-18 months) and is underpinned by two fundamental brain processes involved in perceiving gaze and computing whether gaze is shared between self and other. Therefore the Posner (1980) attentional cueing paradigm was adapted for use with gaze. Friesen and Kingstone (1998) used schematic face stimuli, which depicted a circular face stimulus at the centre of fixation with two open circles for eyes, a small circular nose and a line for a mouth. Initially, the face was presented for 680 ms with the pupils of the face left vacant. Then, closed black circles appeared in the left, right or middle of the eyes for a variable SOA (105, 300, 600 or 1005 ms) after which a target letter (T or F) appeared. This created the depiction of a face either looking directly at the participant or in the same or opposite direction from the target. Participants were required to either classify the location of the letter on the screen (e.g. left or right), identify whether the letter was a T or an F, or simply detect whether a target was present. Participants were aware that the location and identity of the target was random with respect to the validity of the gaze-cue. It was found that, for all but the longest SOAs, when the faces gaze cued towards the target congruently, RTs were faster compared to when the gaze cued away from the target incongruently or when the faces gaze did not cue at all (neutral trials). As such, Friesen and Kingstone (1998) provided initial evidence that gaze could cue attention and that this effect showed hallmarks of being automatic or reflexive in nature. That is, the effect occurred at short SOAs, was relatively short-lived and occurred even when the cues were non-predictive of target location.

At the same time, Driver et al. (1999) conducted three experiments employing a very similar design but using an image of a real face as the cue and with no neutral condition. In their first experiment, the face was presented immediately looking to the left or right hand side of the screen before a variable interval SOA of 100, 300 or 700 ms after which a target letter (T or L) appeared laterally to the face for discrimination. An effect of congruency was present at the 700 ms SOA. The relatively late onset of the cueing effect suggested that the increased complexity of the face and gaze stimulus required more elaborate and time-consuming processing before cueing was initiated compared to the more rudimentary cueing stimuli used in traditional attention research. Therefore, in their second experiment, the face was presented for 900 ms before the gaze-cue appeared to allow for initial face processing to take place before gaze-cueing occurred. In this experiment, gaze-cueing was observed at both the earlier 300 ms and 700 ms SOA. In order to examine the reflexivity of this gaze-cueing effect another experiment was run where participants were informed of a manipulation in which the incongruent cues were four times more likely than the congruent cues. Despite this knowledge, gaze-cueing was still evident at 300 ms, providing further support for reflexivity.

The research by Friesen and Kingstone (1998) and Driver et al. (1999) provided a step change in understanding of gaze processing. Just as these researchers brought attention research into the real world by applying traditional experimental methods to an aspect of social cognition with high ecological validity, further developments saw a shift from the study of perceptual/ attentional aspects of gaze processing to integration with the social-affective domain. Several researchers have pointed out that the reflexive shifts of attention in response to gaze-cues could be employed by the gazer to deceive (Allison, Puce & McCarthy, 2000; Bayliss & Tipper, 2006; Emery, 2000; Klein, Shepherd & Platt, 2009). One example is the use of head feints in basketball, where players gaze in the direction opposite to where they will pass the ball, thereby stalling an interception by the other team. Indeed, the use of head fakes in basketball has now been shown experimentally to produce a cueing effect similar to gaze-cueing where judgements of the direction of a pass are slower when head direction is incongruent with the direction of the pass (Kunde, Skirde & Weigelt, 2011). Bayliss and Tipper (2006) therefore proposed that if faces were consistently observed gazing congruently or incongruently, they would be judged more and less trustworthy, respectively.

To investigate this, a gaze-cueing task was employed using a design in which participants repeatedly observed faces gazing in one of three ways. Some faces were non-predictive of target location, cueing congruently on half of trials and incongruently on the other half of trials. Other faces were predictive-congruent in that they always cued towards the target, or predictive-incongruent, in that they always cued away from targets. On each trial, participants observed a fixation for 600 ms followed by a face gazing toward them for 1500 ms. The faces gaze direction then changed and after a 500 ms SOA, a kitchen or garage target appeared to the left or right hand side of the face. Participants were given the task of either classifying objects as to whether they belonged in a kitchen or garage or were presented to the left or right hand side of the face. After cueing, participants were presented with pairs of faces, one of which was always predictive-congruent and the other always predictive-incongruent. They were asked to choose which face they judged trustworthier, preferred most and believed to be presented most often. The faces in each condition were matched for visual attributes such as age, gender and ethnicity.

In addition to standard gaze-cueing effects, Bayliss and Tipper (2006) found that predictive-congruent faces were more often chosen as trustworthier than predictive-incongruent faces and that this effect negatively correlated with scores on the autism quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin & Clubley, 2001), a measure of the autism phenotype, a developmental disorder characterised by impaired social interactions and communicative ability. There were trends for the predictive-incongruent faces to be preferred less and judged to be presented more often. Therefore, this experiment showed for the first time that the trustworthiness of faces could be learned from identity-contingent gaze-cues. In a subsequent experiment, Bayliss, Griffiths and Tipper (2009) replicated the effects of identity-contingent gaze-cues on trustworthiness judgements whilst at the same time examining the effects of facial expression by using face images from a standardised database (NimStim face database). The effects of gaze cues on trust were only obtained when the face expressed happiness compared to when the faces were neutral or angry. This suggested that effects of gaze cues on trustworthiness judgements were larger when there was a mismatch between the deceptiveness of the behaviour expected by the face and how deceptive the face actually behaved. In this case, because the face is smiling, participants should expect a friendly interaction with the face. However, this expectancy is violated when the face gaze cues incongruently. Therefore these findings suggest the effect might be primarily driven by incongruent faces although the forced choice procedure of both experiments did not allow for identification of the direction of the effect. This effect of enhanced learning when expectancy is violated is consistent with reinforcement learning rules such as the Rescorla-Wagner (1972) model, where the increase in the strength of an association between a conditioned stimulus (e.g. face) and an unconditioned stimulus that evokes emotion (e.g. gaze-cues) is greatest when the difference between what is expected and what happens is maximal, after which learning tails off with diminishing returns after multiple pairings. As the happy faces already predict a positive behaviour, as in congruent cueing, nothing is learned about the trustworthiness of these cues. However, if this were true, one might also expect congruent faces to increase in trust when angry because a negative behaviour is expected but a positive one ensues. The lack of a trust effect for anger shown by Bayliss et al. (2009) may be due to the brain assigning more importance to negative compared to positive stimuli.

More recently, Rogers et al. (2014) have shown that identity-contingent gaze-cueing can also impact upon how favourably the gaze-cuers are behaved towards. Three experiments showed that participants give more money to congruent compared to incongruent faces in economic transactions after identity-contingent gaze-cueing. This occurs regardless of whether the gaze-cueing faces have the possibility of rejecting or returning a fair/ unfair amount proposed by the participant. As such, this suggests that the increased money given to congruent faces is not simply due to the expectation of larger reciprocity and payoffs at a future point in time, but also due to an altruistic sense of social obligation to honour those who share attention. There is evidence that even infants as young as 8 months old can learn the association between identity and contingent congruent and incongruent gaze-cues and use this to guide their attention. Tummeltshammer, Wu, Sobel and Kirkham (2014) found that infants are more likely to overtly follow the gaze of faces that consistently gaze towards locations where interesting objects appear and do not follow the gaze of faces who consistently gaze at locations away from where objects appear. This appears to be specific for faces. The same effects were not obtained when the cueing stimuli were arrows.

In summary, research has shown that identity-contingent gaze-cues affect trust judgements of faces and trust behaviours towards those faces. Much less is known of the processes underlying how perception of gaze and cueing of attention leads to changes in trust judgements. Trust is necessarily a valenced concept with trustworthiness having a positive connotation and untrustworthiness having a negative connotation. Therefore the process whereby faces become charged with positive or negative trust values must involve evaluation. Evaluation is a fundamental component of emotion. It is an elusive representation unlike all others used by the brain. As such, it is formally defined in terms of its effect on behaviour. Also known as reinforcement value, it represents the likelihood that an organism will approach or avoid a stimulus that has biological significance (a primary reinforcer) or which has an association to such a stimulus (secondary reinforcer). Evaluation succeeds sensory processing but precedes motor programming, providing a value signal that can be used to guide behaviour towards or away from particular stimuli. The sophistication of human cognition means we can cognize about these valenced stimuli in a more complex way, adding an extra tier of processing to the system. As of yet, no studies have examined the role of emotion in the learning of face evaluations from identity-contingent gaze-cues. Emotions are largely rapid, implicit processes that are characterised and differentiated by distinct patterns of central and peripheral nervous system activity. The experiments that form the basis of this thesis draw on behavioural, psychophysiological and neuroimaging methods, which can measure both explicit and implicit processes. Before this research is put forward, I will review prior literature, which will put the research described in this thesis into context.

The series of studies described throughout this thesis have used EMG and EEG as implicit measures of emotion. This section will give an overview of these methods to promote a more complete understanding of what each measure represents.

### Electroencephalography (EEG)

The electroencephalogram (EEG) provides a direct measure of neural activity in real time using usually around 16 up to 256 electrodes placed at equal intervals across the scalp. The neural activity that generates the EEG signal is believed to originate from the excitatory and inhibitory post-synaptic potentials generated in the apical dendrites of pyramidal cells in the cerebral cortex. During an excitatory post-synaptic potential (EPSP), the terminal button of a pre-synaptic neuron releases neurotransmitter into the synaptic cleft. This diffuses across to the post-synaptic neuron triggering an influx of positive ions (cations) and an outflux of negative ions (anions) through the cell membrane at the subsynaptic space. The difference in potential between the extracellular space at the subsynaptic region and regions further along the dendrite produces a potential gradient. Because of this potential gradient, cations flow into the subsynaptic space from surrounding regions of the membrane. The negative potential in the extracellular subsynaptic space produces a negative source (sink), which, when coupled with the positive source further along the membrane, creates a current dipole. The opposite occurs when the potential is inhibitory. As EPSPs are a measure of neuronal activation and produce a negative extracellular potential, early researchers began the convention of plotting negative upwards. However, the polarity of the signal recorded on the scalp depends on whether the postsynaptic potentials are excitatory (EPSP) or inhibitory (IPSP) and deep or superficial. Superficial EPSPs and deep IPSPs generate a negative potential whereas deep EPSPs and superficial IPSPs produce a positive potential. Here, deep refers to the membrane further along the dendrites nearer the soma than the apex, which, in pyramidal neurons, is typically close to the surface (Speckmann, Elger & Gorji, 2011).

Around 60,000,000 neurons or a 6cm$^2$ patch of tissue must be synchronously active to summate and produce a measurable signal on the scalp. This is facilitated by the perpendicular arrangement of the dendrites of pyramidal neurons to the cortical surface and the duration of postsynaptic potentials, occurring for up to several tens of milliseconds. When there are periodic bursts of action potentials in a group of synchronised presynaptic neurons, the synaptic activity gives rise to sinusoidal potential fluctuations and the characteristic waves recorded by scalp electrodes. These waves are measured in frequency bands and are believed to be products of different functional networks. These include Delta (<4Hz), Theta (4-7Hz), Alpha (8-15Hz), Beta (16-31 Hz) and Gamma (>32Hz). Due to the convoluted nature of the human brain, the EEG electrodes are primarily sensitive to activity originating from the gyral surface. EEG signals also have to be volume conducted to the electrodes by the resistive surrounding brain tissue, meninges, cerebrospinal fluid, skull and scalp, meaning that the dipoles scalp distribution is low pass spatial filtered. Hence, when the electricity reaches the skull, for example, it spreads laterally shifting the signal from one part of the brain to very distant areas of the scalp. The EEG is not measured in its absolute form. As the voltage recorded on individual scalp electrodes represents the potential for electricity to flow from one place to another, it is scaled as the difference between active electrodes and a reference. Active electrodes are placed across the scalp and are supposed to measure neural activity. The reference electrode is typically placed either on the nose, ears or mastoid bones or is calculated as the average of all scalp electrodes. It is supposed to be as close to the scalp electrodes as possible to pick up and eliminate the same non-brain electromagnetic noise without biasing activity from particular brain sources.

The EEG data can be analysed in one of several different ways. The main distinction between these techniques is in terms of phase-locking. Phase-locked (evoked) activity refers to amplitude modulations that occur at the same time in relation to the onset of a stimulus across all trials. This can be extracted by simply averaging across multiple trials in a certain condition to produce the event-related potential (ERP). In contrast, non-phase locked (induced) activity, is primarily concerned with measuring oscillations associated with a particular stimulus. Because the phase of oscillatory activity varies across trials, a simple averaging procedure would mean that the oscillations would cancel each other out producing no measurable signal. Therefore, oscillatory power in different frequency bands are quantified using non-phase locked statistics in each trial before being averaged. This can be used to show event-related synchronisation/ desynchronisation (Pfurtscheller & Lopes da Silva, 1999). There are two theories of the generators underlying ERPs. Amplitude-modulation theory states that the stimulus evokes a transient increase in amplitude of the EEG that is time locked to stimulus onset. In contrast, the phase-resetting view supposes that the stimulus shifts the phase of on-going EEG oscillations so that they become aligned for a period of time following stimulus onset. The amplitude of the oscillation is the same both pre- and post-stimulus but the alignment of the phase of the EEG oscillation means that the oscillations sum up to produce an observable ERP. Evidence for and against either mechanism has been shown. Therefore, the shared generator hypothesis proposes that both mechanisms are at play. This theory states that the ERP may be the product of dormant cells becoming active in response to a stimulus at the same time as on-going oscillations entrain their phase to the stimulus (Bastiaansen, Mazaheri & Jensen, 2012).

### Facial Electromyography (fEMG)

Facial electromyography (fEMG) measures action potentials associated with the contraction of facial muscles during emotional expression. The face consists of 43 distinct muscles. Each muscle is made up of many muscle fibres that are surrounded by an excitable cell membrane called the sarcolemma. Muscle fibres are in turn made up of many myofibrils. Myofibrils are composed of overlapping myosin (thick) and actin (thin) filaments and divided into segments called sarcomere. Myosin filaments have a series of protrusive globules, known as heads, that hold a molecule of adenosine triphosphate (ATP), a source of energy. Both filaments are negatively charged. The myofibrils are surrounded by an intracellular sac, the sarcoplasmic reticulum, which contains positive calcium ions. When acetylcholine (Ach) is released by a lower motor neuron into the neuro-muscular junction, an action potential sweeps along the sarcolemma through a network of transverse tubules to the sarcoplasmic reticulum, triggering the release of calcium ions into the space where myosin and actin are located. When the calcium ions bond to the actin filament, this attracts the myosin, pushing its head further along the actin, expending the ATP. This forces the actin and myosin to move past each other. The release of energy by ATP, frees the calcium from the actin filament and pumps it back to the sarcoplasmic reticulum. At the same time, the actin and myosin separate. The globular head on the myosin filament is then restocked with ATP. As this process repeats, the actin is ratcheted up along the myosin, the muscle contracts and a facial expression occurs (Tassinary, Cacioppo & Vanman, 2007).

Single neurons synapse onto multiple muscle fibres and are referred to as motor units. The ratio of motor units to muscle fibres determines the strength of actions or whether they are fine and detailed. The facial muscles have a ratio of as few as 3 muscle fibres per neuron allowing for detailed patterning of expressions. In comparison, the gastrocnemius muscle in the leg has around 2000 muscle fibres per neuron. The fEMG signal represents the summation of many motor unit action potentials (MUAP) within the same muscle, which are called motor unit pools. This is measured with a bipolar electrode montage and differential amplification. Two electrodes are placed along the muscle 1-2 cm apart. When MUAPs are triggered, the wave of depolarisation spreads along the muscle, reaching one electrode before the other, producing a dipole. A difference in potential between electrodes is measured and there is an inflection in the fEMG, which reverses in polarity when the depolarization zone comes closer to one or the other electrodes. To control for electromagnetic noise, before differential amplification the signal from a reference electrode placed on the forehead is subtracted from both electrodes to eliminate noise signals that will be common to both.

Lower motor neurons innervating the facial muscles are part of the seventh cranial nerves originating from the facial nucleus of the caudal pontine tegmentum in the brain stem. They are controlled by upper motor neurons descending from multiple brain areas including the primary motor cortex (M1), ventrolateral premotor cortex (PMCvl), supplementary motor area (SMA) and anterior (M3) and caudal (M4) face areas of the midcingulate cortex. M1 and PMCvl innervate the lateral segment of the contralateral facial nucleus and control the lower facial muscles. The SMA and M3 innervate the motor neurons in the medial segment of the facial nucleus and control upper facial muscles. M4 innervates the lateral facial nucleus, in particular, the upper lip. M1, PMCvl and SMA are involved in the voluntary control of facial movements. Patients with lesions to these areas caused by cerebrovascular accident (CVA) are impaired in the voluntary generation of a symmetrical smile, yet when presented with jokes to induce humour can smile normally (Hopf, Muller-Forell & Hopf, 1992). In contrast, CVAs in the vicinity of the midcingulate (M3 and M4) area have been shown to leave voluntary facial movements intact whilst impairing spontaneous emotional expressions

(Karnosh, 1945). The amygdala has been proposed to play a role in the decision to facially express particular emotions in particular contexts by way of its projections to M3 and role in evaluating social stimuli such as others' facial expressions, identity and gaze (Gothard, 2014). Both amygdala and M3 receive and project to the viscera and insula and may monitor the sensory consequences of facial expressions to continually update them in a way that is appropriate for the current context (Livneh, Resnik, Shohat & Paz, 2012).

The facial motor nuclei are proximal to autonomic control centres suggesting similarity and coupling between facial expressions and visceral emotional arousal states (Porges, 1997). Interestingly, relative neocortex size (Shultz & Dunbar, 2007), facial motor nucleus volume (Dobson & Sherwood, 2011) and range of facial movements (Dobson, 2009) are all correlated with group size across catarrhines. Furthermore, facial motor nucleus volume is correlated with neocortex size (Dobson, 2012). This has led Dobson (2012) to propose that the neocortical expansion underlying selection for enhanced executive brain functions results in greater facial expressivity due to concurrent expansion of the strongly connected facial motor nucleus. As highly expressive individuals are more likely to cooperate (Mehu, Grammer & Dunbar, 2007; Schug, Matsumoto, Horita, Yamagishi & Bonnet, 2010) and be cooperated with (Krumhuber, Manstead, Cosker, Marshall, Rosin & Kappas, 2007; Stouten & De Cremer, 2010; Wubben, De Cremer & van Dijk, 2011), the increased sociality of catarrhine species is believed to be due to this increased expressivity providing the opportunity to better communicate and read the trustworthiness of potential social partners.

I will now review research that has drawn on neuroimaging and psychophysiological methods to shed light on issues that could contribute to our understanding of the psychological and neural processes underlying the role of emotion in learning face evaluations from gaze-cueing.

## Face Processing

Faces have a significance to us like no other visual stimulus and rightly so. The human face has multiple channels of information of high value to social interactions. These include information about identity, expression of emotion, gender, race, attractiveness, speech and direction of attention. Some of these aspects of the face appear to be processed independently, at least in the early stages of face processing (Bruce & Young, 1986). Much of what is known about face perception has been informed by neuroimaging data which has found distinct regions of the brain that appear to be activated by different facial attributes.

These findings can be understood within the framework proposed in Haxby, Hoffman and Gobbini's (2000; 2002) human neural system for face perception. According to this model, a core system consists of three interlinked regions. The inferior occipital gyrus (occipital face area; OFA) is involved in the early visual processing of the face and this visual representation feeds forward to the superior temporal sulcus (STS) and lateral fusiform gyrus (fusiform face area; FFA) which process the dynamic aspects of the face (gaze, expression, lip movement) and invariant aspects of the face necessary for perception of identity, respectively. These nodes provide representations of the face that can be used by further neural regions in an extended system. The representation of gaze direction can trigger spatial attention in a congruent direction via the connection between STS and intraparietal sulcus (IPS) and the representation of expression (in STS) and identity (in FFA) can trigger emotional reactions via associations with various regions involved in emotion processing including the amygdala, insula and orbitofrontal cortex. The representation of invariant facial features by the FFA is used by anterior temporal (ATL) areas for recognition of identity.

Face processing has also been studied with ERPs. There appears to be three ERP components that are particularly involved in face processing, the P100, N170/ VPP and N250. The P100 has an occipital distribution and is believed to reflect low level visual processing. Some studies have shown it is more responsive to faces than objects (Dering, Martin, Moro, Pegna & Thierry, 2011; Herrmann, Ehlis, Ellgring & Fallgatter, 2005) whereas others have failed (Boutsen, Humphreys, Praamstra & Warbrick, 2006; Rossion, Joyce, Cottrell & Tarr, 2003). The P1 difference between faces and non-faces is most likely due to the low level visual properties of the stimuli. When these are strictly controlled for colour, luminance, size and spatial frequency, face selectivity is abolished (Rousselet, Husk, Bennett & Sekuler, 2005). Activity in the P1 range appears to be important for subsequent face processing. TMS delivered to the OFA at 60 ms and 100 ms enhances subsequent ERP amplitude in the N170/ VPP range (Sadeh, Pitcher, Brandman, Elsen, Thaler & Yovel, 2011).

The N170 and VPP are generally regarded to underlie the structural encoding of faces. Both the N170 and VPP reflect the same brain activity occurring in the same time range but as a positivity on mid-central electrodes for the latter and as a negativity on posterior occipitotemporal electrodes for the former (Joyce & Rossion, 2005). Therefore the N170 and the VPP reflect opposite ends of the same electrical dipole in a similar way to which the other visual components responsive to faces are inverted on anterior compared to posterior electrodes. The N170 and VPP have consistently been shown to be enhanced for faces compared to non-faces (Bentin, McCarthy, Perez, Puce & Allison, 1996; Jeffreys, 1996). The components are regarded as the time points at which basic-level categorization of the face as a face occurs. For example, when the categorization of the face is made more difficult by inversion (Bentin et al., 1996; Rossion et al., 1999; 2000) or offsetting half of the face (Letourneau & Mitchell, 2008), the N170 is delayed. However, it is less clear whether the N170/VPP is involved in face identification. Face identity representation is tested with repetition paradigms where suppression of responses to test faces succeeding faces of the same identity is believed to show when identity is represented. Two different types of repetition have been studied- immediate and delayed. Immediate repetition is the same as the adaptation paradigm where test and adaptor are presented in rapid succession. Using this paradigm, some studies have observed repetition suppression of the N170 (Campanella et al., 2000; Guillaume & Tiberghien, 2001; Itier & Taylor 2002), whereas others did not (Huddy, Schweinberger, Jentzsch & Burton, 2003; Schweinberger, Huddy & Burton, 2004; Schweinberger, Pickering, Jentzsch, Burton & Kaufmann, 2002). However, the positive findings could simply be due to repetition of visual attributes as opposed to identity. More recently, it was shown that the identity repetition effect in N170 remained despite use of different photographs and viewpoints across test and adaptor (Caharel, d'Arripe, Ramon, Jacques & Rossion, 2009, Jacques, dArripe & Rossion, 2007). However, the N170 has also been shown to be reduced when preceded by another face regardless of whether it is of the same or different identity (Kloth, Schweinberger & Kovacs, 2010; Maurer, Rossion & McCandliss, 2008). Thus, the N170 has not been unequivocally implicated in representing identity. In contrast, the N250 is more strongly modulated by repetition.

The N250 is a negative deflection at 250 ms measured on posterior occipitotemporal electrodes. It is not typically measured as a stand-alone component but is identified in specific paradigms. It was first identified as a potential marker of face recognition in the immediate repetition paradigm. This is known as the N250r- for repetition. It is characterised by a larger negative deflection for a face that is preceded by a face that is of the same identity compared to when the preceding face is not of the same identity (Schweinberger et al., 2002; 2004). The N250r is not affected by distortions such as stretching of the prime stimulus, suggesting that it is not driven solely by low-level visual attributes and represents the more abstract process of recognition (Bindemann, Burton, Leuthold & Schweinberger, 2008). The N250 has also been studied in paradigms where trials on which the faces were unfamiliar to participants are compared to trials where the faces have become familiar after multiple exposures (Pierce, Scott, Boddington, Droucker, Curran & Tanaka, 2011; Tanaka, Curran, Porterfield & Collins, 2006). This paradigm also shows an N250 that is larger when faces are familiar. Interestingly, both immediate repetition and familiarisation procedures have been integrated into the paradigm to study view-independent face recognition (Zimmerman & Eimer, 2013). This study showed that the N250r to immediately repeated faces was evident when both prime and test faces were of the same orientation in both early and later blocks. However, only in later blocks, after faces had become familiar from a variety of viewpoints, was the N250r evident for repeated identities that differed in viewpoint.

Representation of identity allows faces to be stored in memory, recognised, differentiated from each other and can prime knowledge and emotions. It relies upon coding subtle differences in the unique shape and spatial relations of the facial features. This is known as holistic or configural processing. It is illustrated by the effects of inversion and chimera. When faces are inverted, they are recognised more poorly compared to inverted objects (Yin, 1969) and when the top half of one face is put with the bottom half of another face, participants are poorer at identifying the faces from which the parts are drawn (Young, Hellawell & Hay, 1987). Whereas inversion interferes with the process of binding features together holistically, the fusion of the parts of the chimeral face into a unique whole interferes with breaking down the face into its constituents. At its most simplest, face recognition proceeds from a view-dependent visual description of the structure of the face to a view-invariant description that can be matched to face representations stored in memory, trigger feelings of familiarity and allow the retrieval of names (Bruce & Young, 1986).

Valentine (1991) has proposed that faces are represented and distinguished by their deviation along the multiple dimensions upon which faces vary. This is known as face space. The nature of the face representations in memory have been investigated with visual adaptation (Leopold, O’Toole, Vetter & Blanz, 2001; Rhodes & Jeffery, 2006). Adaptation designs involve assessing the effects of observing one stimulus (the adaptor) on a subsequently presented test stimulus, when these stimuli vary upon a particular stimulus dimension (Frisby, 1979). The logic behind the technique is that, if two stimuli are coded along the same dimension, then adaptation in cell populations coding the adaptor stimulus should bias the perception of the test stimulus along that dimension in the opposite direction to the adaptor. By morphing an individual face (Jim) with a face compiled from the average of multiple faces it is possible to create a trajectory in facial appearance varying from the original face (Jim) to a face that has an appearance that is opposite to the original face in face space (anti-Jim). When participants adapted to anti-Jim, average faces were more likely to be classified as Jim than when there was no adaptation or when the face was not opposite from

Jim in face space. These findings support a norm-based identity-coding model whereby faces are represented in terms of their difference from the average on multiple face dimensions.

The anterior temporal lobes (ATLs) appear to be the most likely candidate regions for face recognition (Anzellotti & Caramazza, 2014; Collins & Olson, 2014). Evidence comes from neuropsychology, neuroimaging and neurophysiology. Prosopagnosia, a disorder characterised by a marked deficit in the recognition of faces, typically occurs with damage to the ATLs or the pathways to the ATLs (Damasio, Tranel & Damasio, 1990; Ellis, Young & Critchley, 1989; Evans, Heggs, Antoun & Hodges, 1995; Tranel, Damasio & Damasio, 1997). FMRI has shown activation in patches of the ATL in response to faces (Tsao, Moeller & Friewald, 2008; Rajimehr, Young & Tootell, 2009) that correlates with behavioural recognition performance (Nasr & Tootell, 2012) and can be used to differentiate faces with multi-voxel pattern analysis (MVPA; Kriegeskorte, Formasino, Sorger & Goebel, 2007). Single-unit recording in macaques has shown that individual neurons in ATL are selective for different faces (Freiwald & Tsao, 2010). This is despite the fact that all the faces were unfamiliar to the monkeys suggesting that the representation in ATL is not exemplar based or multichannel coded, but rather, faces unique appearances are represented on the dimensions along which faces previously encountered vary. Indeed, the activity of these cells increases linearly with the difference of a face from the average (Leopold, Bondar & Giese, 2006), suggesting a neural correlate of norm-based identity coding.

Face selectivity has also been recorded with intracranial electrodes in the ATL of humans, peaking at around 350 ms (Allison, Puce, Spencer & McCarthy, 1999). Interestingly, like the single cells recorded in the ATL of monkeys, the N250 component recorded in humans is modulated by face identity strength, which is defined as the extent of the difference of a faces appearance from an average face, suggesting that the activity of these cells may be the underlying source of the N250 (Zheng, Mondloch, Segalowitz, 2012). In contrast to the ATL, the FFA has not been consistently shown to be sensitive to identity. The FFA is face selective, responding more to faces than objects (Kanwisher, Mcdermott and Chun, 1997). However, it appears to be more involved in processing the perceptual attributes necessary for identification than in memorising faces. FFA activation is larger for intact faces than for scrambled faces (Liu, Harris & Kanwisher, 2009) and there is more release from adaptation when there are changes to internal or external features (Andrews, Davies-Thompson, Kingstone & Young, 2010), when the spatial relations between internal features of an individual face are altered (Rhodes, Michie, Hughes & Byatt, 2009), when different images of the same person are presented, regardless of whether they are familiar or unfamiliar (Davies-Thompson, Gouws & Andrews, 2009) or when the physical difference between adapt and test faces crosses an identity boundary, where identification of the face shifts from one person to another (but this does not correlate with familiarity of the face unlike in ATL; Rotshtein, Henson, Treves, Driver & Dolan, 2005). MVPA in the FFA also cannot discriminate individual faces like in the ATL (Kriegeskorte et al., 2007). Thus, while the FFA is involved in face processing necessary for face representation and recognition it is likely that this processing provides an input to the ATL where recognition takes place.

## Gaze Processing

High visual acuity costs the brain space and energy. Consequently, the restricted visual field size of the fovea has to be constantly moved via attention and oculomotor systems for a detailed representation of objects in different spatial locations to be maintained. Such refixations are readily perceivable by others aided by the high contrast between the human iris and sclera (Kobayashi & Kohshima, 1997; 2001) providing valuable visual cues to objects of motivational significance and likely sources of action. As such, eye-gaze direction detection features prominently in models of social cognition and appears to be the most important cue to "social attention" (Baron-Cohen, 1995; Emery, 2000; Langton, Watt & Bruce, 2000). Indeed, experimental evidence shows humans are adept at fine perceptual discrimination of gaze direction (Anderson, Risko & Kingstone, 2011; Gibson & Pick, 1963) and at shifting attention to detect objects in the line of others' sight (Langton, O'Donnell, Riby & Ballantyne, 2006). Gaze appears to be represented by a multi-channel coding system, whereby the detailed representation of gaze direction is encoded in the relative activity of cells that code leftward, rightward and direct gaze. Evidence comes from the visual adaptation paradigm. Adaptation to gaze oriented 25° to the left biases faces with gaze oriented 5° or 10° to the left to be judged as gazing directly (Jenkins, Beaver & Calder, 2006). Furthermore, adaptation to direct gaze leads to 5° and 10° left and rightward gaze to be categorised as non-direct, adaptation is more pronounced for stimuli on the adapted side and adaptation to leftward and rightward gaze concurrently still produces adaptation aftereffects for less extreme test stimuli (Calder, Jenkins, Cassell & Clifford, 2008).

Gaze processing is highly canalized, being present early in infancy, as shown by the tendency to follow others' gaze (Farroni, Massaccesi, Pividori, Simion & Johnson, 2004; Hood, Willen & Driver, 1998), and is important to social development (Striano & Reid, 2006). Establishing joint/ shared attention appears to be particularly important. Joint attention is defined as when two individuals attend to the same object. Shared attention requires that the individuals concerned are aware that the other is attending to the same object (Emery, 2000). This is known as a triadic representation as it requires the knowledge of the relations between three entities: the other person, the self and the object commanding attention. Baron-Cohen (1995) recognises the importance of gaze in his proposal of four modules that underpin social cognition. These include an intentionality detector (ID), eye direction detector (EDD), shared attention mechanism (SAM) and theory of mind module (TOMM). All of these nodes are proposed to work in tandem. The ID detects the intentionality of the gaze shift, the EDD detects gaze direction, interprets this as "seeing" and feeds-forward to the SAM, which computes whether attention is shared between self and other. The highest level of social understanding, TOMM, uses inputs from the SAM to form representations of others epistemic mental states and how these mental states relate to actions. This is facilitated by the ability of SAM to form triadic representations, in this case to understand the relations between an individual, their mental state and the state of affairs of the world.

Multiple cues can be used to ascertain what another person is attending to. These cues include head, body and gaze direction. Unsurprisingly, eye-gaze direction is the most important cue to the direction of another persons visual attention. This is because we perceive the environment with our eyes, which is the most dominant sense in most humans. Although head and body orientation are usually correlated with gaze direction, only gaze direction is guaranteed to accurately show where a person is attending. However, under restricted viewing conditions in which the eyes are not visible, people may have to rely upon the only other visual cues that are available. Head orientation is the second best cue to visual attention due to the tendency to move the head in the direction of looking. This is less useful than gaze direction because gaze direction can vary in the eye sockets. The head can point one way but gaze can be directed in the opposite direction. If head orientation is not visible, we must rely upon body orientation as it is usually correlated with gaze direction due to the tendency to face and walk in the direction of looking. However, this is less useful than gaze and head direction because gaze direction can vary in the head and head direction can vary about the body. This gradation in the importance of different cues to social attention was first demonstrated in the response properties of cells in the anterior superior temporal sulcus (aSTS). These cells have been shown to be responsive to the gaze, head and body orientation of others to compute attention direction (Morin, Hadj-Bouziane, Stokes, Ungerleider & Bell, 2014; Perrett, Heitenan, Oram & Benson, 1992). Importantly, cells that are selective for a particular head orientation can be overridden by gaze direction and cells responsive to a particular body orientation can be overridden by head orientation (Perrett et al., 1992).

Evidence for such cells in humans has been demonstrated with fMRI. The aSTS shows greater repetition suppression to repeated faces that have the same gaze direction (Calder et al., 2007) and MVPA has shown that voxels in the aSTS code gaze direction whilst controlling for head orientation (Carlin, Calder, Kriegeskorte, Nili & Rowe, 2011).

However, gaze perception per se is not particularly useful. The remarkable ability of humans to perceive small gradations in others' gaze direction is due to the value of detecting objects in the line of others sight. This requires the additional step that attention is accurately cued to the most likely object attended to by the gazer. The rapidity and automaticity of this process is illustrated in studies of gaze-cueing (Driver et al., 1999; Friesen & Kingstone, 1998; Frischen, Bayliss & Tipper, 2007). Gaze-cueing is not affected by the knowledge that the target is more likely to appear in the opposite direction of the gaze (Driver et al., 1999), not reduced by visual and verbal working memory load (Law, Langton & Logie, 2010), can occur under high perceptual load conditions, such as in rapid serial visual presentation, and without awareness, under flash suppression and backward masking (Sato, Okada & Toichi, 2007; Xu, Zhang & Geng, 2011). Gaze-cueing appears to be driven specifically by perception of the eyes rather than low-level directional information. Visual adaptation to a gazing face reduces the gaze-cueing effect when the adaptor face gazes in the same direction as the cueing face, but not when the adaptor is an arrow stimulus (Bayliss, Bartlett, Naughtin & Kritikos, 2011). Gaze-cueing effects are not limited to still images of faces presented in computerised experiments. Lachat, Conty, Hugueville and George (2012a) have also demonstrated gaze-cueing effects in a novel paradigm in which gaze-cueing was observed between participants engaged in a face-to-face interaction with an experimenter.

The intraparietal sulcus (IPS) appears to play a role in the attention shift by gaze. The IPS is associated with the dorsal frontal-parietal attention network (Corbetta, Patel & Shulman, 2008; Corbetta & Shulman, 2002). Neurons within a specific region of the IPS, termed the lateral intraparietal area (LIP), code the saliency and worthiness of the stimulus to be attended to with a saccade (Bisley & Goldberg, 2010; Colby & Goldberg, 1999). These neurons have retinotopic selectivity, responding to stimuli presented at a specific point in space. The responses are enhanced by behavioural relevance, being larger when a saccade is to be made to the stimulus, when a behavioural response is cued by the stimulus (Colby & Goldberg, 1999) or when the stimulus is a reward (McCoy & Platt, 2005). LIP responses are also modulated by social target value being larger in amplitude to social stimuli which are of particular importance to the monkey, such as the face of a more dominant monkey, as shown by behavioural measures (Klein, Deaner & Platt, 2008). Most importantly though, research has shown that LIP neuron activity is enhanced when a monkey face is observed gazing in the same direction as the visual field of a to be presented target stimulus compared to when the face gazes away from the targets visual field (Shepherd, Klein, Deaner & Platt, 2009). The time-course of this neural activity occurs within the latency of the effect of gaze-cueing on reaction times measured from eye-movements (Deaner & Platt, 2003).

These findings could be interpreted in of two ways. Neurons in the ventral premotor cortex (vPMC) and the inferior parietal lobule (IPL) of macaques have been shown be involved in both executing specific hand and facial actions but are also activated by the sight or sound of the same actions (di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolatti, 1992; Fogassi et al., 2005; Gallese, Fadiga, Fogassi & Rizzolatti, 1996; Kohler et al., 2002) and neuroimaging provides evidence for the existence of similar responses in similar regions in humans (Caspers, Zilles, Laird & Eickhoff, 2010). The similarity of responses to observed and executed actions is widely believed to be indicative of a “mirror system”, in which observers use their own motor repertoire to simulate what it would be like to execute the action themselves to understand others’ actions and intentions. Therefore the activity in LIP neurons, which represents the importance of a stimulus location to be gazed towards, and which is larger when cued by a face gazing congruently to that location, could reflect mirroring of attentional shifts. However, although the attention shift is by the gaze of another monkeys face, these LIP responses could simply reflect domain-general attentional orienting. It is difficult to decide between these two alternatives. The optimal control experiment would be to use non-social stimuli such as arrows but such symbolic stimuli are not recognised as directional cues as they are in humans who are acculturated to such stimuli through learning.

Another region that is important for gaze-cueing is the posterior STS. When muscimol, a GABA agonist that inhibits neural activity, is injected into the left pSTS region during a gaze-cueing task, gaze-cueing is abolished (Roy, Shepherd & Platt, 2012). Thus, it has been proposed that pSTS may feed gaze information into IPS for orienting. In fMRI studies, pSTS activity is also larger during observation of incongruent compared to congruent gaze shifts (Pelphrey, Morris & McCarthy, 2005; Pelphrey, Singerman, Allison & McCarthy, 2003). This may be due to extra processing resources needed to understand the intentions of a person who chooses to gaze away from a highly salient and interesting object in the absence of any competing stimuli. Indeed, the pSTS is activated by geometric shapes, which move in ways that are perceived to have intentionality (Castelli, Happe, Frith & Frith, 2000).

The temporal dynamics of gaze processing have been studied using ERPs. The N170 responds equally, if not stronger, to eyes presented in isolation as it does to whole faces

(Bentin et al., 1996; Itier, Latinus & Taylor, 2006). In contrast, other face parts such as the nose or the mouth do not elicit as strong response (Bentin et al., 1996). Thus the eyes appear to be particularly important for early face processing. Studies have also examined ERPs during gaze-cueing. These studies have shown that the P100 and N170 to laterally presented target stimuli are amplified when a centrally presented face gazes toward the target, congruently, compared to when gazing away, incongruently (Schuller & Rossion, 2001; 2004; 2005). In addition, the P300 component, a positivity on posterior electrodes at around 300 ms, was found to be larger on incongruent compared to congruent trials. This effect of congruency on P300 amplitude is also present when the target is presented before the gaze-cue and disappears at the onset of the cue (Senju, Johnson & Csibra, 2006; Tipples, Johnston & Mayes, 2013). This suggests that it does not simply reflect gaze orienting effects but rather the violation of expectancy and possibly increased processing of intentions, similar to the fMRI studies implicating the pSTS in this process. However, the same effects were also observed for arrow cues suggesting that this is a domain general neurocognitive process although the low spatial resolution of EEG leaves open the possibility that the effects for arrows and gaze are mediated by adjacent but separate areas of the brain. Another study has examined oscillatory activity during live joint attention. This study showed that alpha (8-12 Hz) power on left central-occipito-parietal electrodes was reduced during joint attention compared to non-joint attention (Lachat, Hugueville, Lemarechal, Conty & George, 2012b). The alpha rhythm is typically reduced under high attentional demands and arousal. This reflects increased cortical processing and could be similar to mirroring of attention in LIP.

## Emotion

Emotions are states elicited by the experience, omission or termination of rewards and punishers. Emotions are both valent (positive or negative) and arousing (activating). Emotions are the way genes, refined by natural selection, specify the goals for actions, which increase the fitness of the organism (Rolls, 2008). Rewards motivate approach behaviours to stimuli that sustain life and punishers motivate avoidance behaviours away from stimuli that threaten life. Emotions are characterised by changes in cognition, behaviour and physiology. Cognitive processes that are involved in emotion include the initial perceptual processing of stimuli that elicit the emotion, the appraisal of stimuli as rewards and punishers, attentional prioritisation of emotional stimuli, value based reasoning, decision-making and memory storage and retrieval of emotional events. Behavioural processes include approach and avoidance or generally the actions that are contingent upon the emotional stimulus. Higher order areas of the brain in the pre-frontal cortex typically plan and execute these behaviours, although, after learning, the acts may become more automated. Physiological activity encompasses changes in visceral, facial motor, respiratory, genital, endocrine and immune system activity that prepares the body to process, consume, reject, confront or act on the emotional stimulus. These changes are rapid and automatic, outside of conscious control.

Facial expressions appear to be a particularly special type of emotional expression in that these responses appear to also have a communicative function that regulates social interactions, at least in humans and some non-human primates. Research has shown that the facial expressions of fear, anger, disgust, happiness, sadness and surprise can be communicated between vastly different and unrelated cultures and therefore may be universal and highly canalized (Ekman & Friesen, 1971). The history of theory in emotion research has been one of repeated debate about the role of physiological activity/ feedback in generating emotional states. James (1884) and Lange (1885) originally proposed that different patterns of physiological signals arising from the peripheral nervous system were used by the brain to infer different emotions. One of the downfalls of this theory was that the stimulus must be appraised as emotional before physiological changes can occur in the first place. In addition, physiological activity during different emotion states are too similar to produce the whole variety of different emotions that one can experience. The importance of appraisal is demonstrated by Schachter and Singer (1962) who showed that administration of adrenaline, which elicits physiological arousal, elicited stronger emotions in those individuals who could not attribute their physiological sensations elicited by the drug to the drug. Therefore, subsequent theories have emphasised the role of appraisal in the generation of emotions (Frijda, 1986; Izard, 1993; Lazarus, 1991; Scherer, 1999). This can be operationally defined as the evaluation of an event as relevant to a goal that is important, eliciting positive emotion when an event advances that goal and negative emotion when the event impedes the goal.

More recently, interest in the contribution of physiological signals to emotion has been reignited in the form of the somatic marker theory of emotion (Damasio, 1996). According to this theory, somatic states (musculoskeletal, visceral) are triggered by stimuli associated with the emotion in question by the ventromedial sector of the prefrontal cortex. These somatic states are then perceived in the somatosensory cortices and insulae and in combination with the knowledge of what triggered the somatic marker in the first place, are then interpreted or marked as positive or negative either consciously or unconsciously and used for decision-making. The somatic markers can also bypass the activation of the peripheral nervous system by sending the somatic marker directly to the somatosensory cortex. In this case, the somatic states are imagined or simulated rather than experienced directly. Supportive evidence was provided by Bechara, Damasio, Tranel and Damasio (1997). This study used the Iowa gambling task in which participants have to choose cards from one of four different card decks. Card decks A and B lead to less gain in the short term, giving $50 on each turn whereas decks C and D lead to more gain in the short-term, with a value of $100 but which, every now an then, also lead to a penalty, which in the long-term leads to a net loss in money. The optimal strategy in this task is to learn which decks lead to net losses and which lead to net gains over repeated trials. Bechara et al. discovered that normal participants showed a behavioural aversion to the disadvantageous card decks and an enhanced skin conductance response before choosing from these decks. In contrast, patients with ventromedial prefrontal cortex lesions showed no aversion to the disadvantageous decks and no anticipatory skin conductance response when choosing from those decks.

The notion that embodied emotional states influence decision making is consistent with recent developments in embodied cognition theory. According to embodied cognition theories, concepts and the cognitive operations that take them as inputs are underpinned by simulations in sensory-motor cortices (Barsalou, 1999; Niedenthal, 2007). This theory provides a realistic way in which cognition is represented in the brain and how it could have developed from more basic sensory-motor functions. There is now a wealth of data supporting this type of representation by the brain. One example is the modality switching cost effect, where reaction times to verify perceptual properties of an object are quicker when preceded by a trial on which the properties verified were in the same modality compared to when in a different modality (Pecher, Zeelenberg & Barsalou, 2003). There is also the action compatibility effect, where sensibility judgements of sentences describing actions are quicker when the direction of the response is in the same direction as the action described in the sentence (Glenberg & Kaschak, 2002). There is also neuroimaging support. For example, perceiving and understanding others' actions recruits the motor cortices (di Pellegrino et al., 1992), as does understanding of action related words (Hauk, Johnsrude & Pulvermuller, 2004; Pulvermuller, 2005) whereas observing others being touched activates somatosensory areas (Keysers et al., 2004) and words representing concepts associated with perceptual processing in different sensory modalities including in the visual, auditory, haptic, gustatory, and olfactory domains, activate corresponding sensory areas of the brain (Gonzalez et al., 2006; Kiefer, Sim, Herrnberger, Grothe & Hoenig, 2008; Simmons et al., 2007).

The stimuli that evoke emotions by themselves, without the need for learning, are referred to as primary or unconditioned reinforcers. However, stimuli that become associated with primary reinforcers through repeated pairing can also take on the affective properties of the unconditioned reinforcer through the learning process known as classical conditioning and become known as secondary reinforcers. This has typically been studied in animals. For example, Pavlov (1927) first showed in dogs that after repeated associations between a bell (conditioned stimulus or CS) and the subsequent delivery of food (unconditioned stimulus, US), the bell itself could elicit the same (conditioned, CR) response, salivating (unconditioned response, UR), as the US itself. Similarly, if a tone presented to rats is repeatedly associated with the subsequent administration of an electric shock to its feet it typically shows a range of fear responses to the tone itself including freezing, autonomic arousal, hypoalgesia, reflex potentiation and stress hormone release (Ledoux, 2000). Classical conditioning is widely regarded to reflect a learning process that allows organisms to predict and prepare for an encounter with motivationally important stimuli (Konorski, 1967). In contrast to animals, humans have the ability to verbally report how they feel and how much they like particular stimuli. Research examining ratings and choice of stimuli that have been associated using classical conditioning procedures in humans is typically referred to as evaluative conditioning (Hofmann, De Houwer, Perugini, Baeyens & Crombez, 2010). Typically, when neutral CSs are associated with liked or disliked stimuli, which act as USs, the CSs become liked or disliked, respectively. Evaluative conditioning effects have been obtained with a variety of different stimuli from a variety of different sensory modalities acting as the CSs and the USs (De Houwer, Thomas & Baeyens, 2001). One example is the finding of increased and decreased liking of faces which are associated with liked or disliked odors, respectively (Todrank, Byrnes, Wrzesniewski & Rozin, 1995). Despite the similarity between the classical conditioning procedures used with animals and the evaluative conditioning procedures used with humans, there is no consensus as to whether the two types of learning are underpinned by the same associative learning rules (Hofmann et al., 2010).

There are several models of classical conditioning. Rescorla and Wagner (1972) proposed that association learning occurs when there is an error between what a CS predicts and what US actually occurs. Learning is greater when there is a larger mismatch between what is expected based on all CSs that are present and what USs actually occur. Other theories also exist that are based on attention. Mackintosh (1975) and Pearce and Hall (1980) propose that learning is mediated by increased attention to CSs that are more reliable predictors of the USs. In contrast to traditional associative learning theories based on animal data, theories of evaluative conditioning have developed to explain some apparent inconsistencies between classical and evaluative conditioning. In particular, evaluative conditioning has been found to be insensitive to extinction, cue competition (such as blocking; Beckers, De Vicq & Baeyens, 2010; Laane, Aru & Dickinson, 2010) and to occur both with and without awareness of the contingencies between CS and US (De Houwer et al., 2001). These theories include the referential, holistic, implicit misattribution, conceptual categorization and propositional accounts. According to the referential account, the CS activates a representation of the US via referential associations but not a prediction that it will occur (Baeyens, Eelen, Crombez & Van den Bergh, 1992). Learning is due to the co-occurrence of stimuli rather than contingency of one stimulus upon another. The holistic account states that, due to the co-occurrence of the CS and the US, both stimuli are integrated into the same representation (Martin & Levey, 1978; 1994). The value of one stimulus therefore generalises to the other. The implicit misattribution account states that the affective response to the US is misattributed to the CS (Jones, Fazio & Olson, 2009). The conceptual categorization account states that the pairing of the CS with the US allows participants to notice similarities of the liked/ disliked features of the US that are present in the CS (Davey, 1994). Participants then categorise the stimulus as liked or disliked based on this similarity. Finally, the propositional account states that evaluative conditioning depends upon the non-automatic formation of abstract propositional knowledge about the relations between CS and US (De Houwer, 2006). Participants then justify liking of the CS based on the US.

Another way in which emotion may be elicited and associated with stimuli through the process of evaluative learning is through the elicitation of fluency. Fluency refers to the ease with which a stimulus is processed perceptually or responded to. Perceptual fluency has been manipulated using priming, protypicality or stimulus presentation duration. It has been shown that stimuli which are primed, more prototypical or presented for a longer duration are typically rated more positively than those which are not primed, less prototypical or presented for a shorter duration (Winkielman & Cacioppo, 2001; Winkielman, Halberstadt, Fazendiero & Catty, 2006). These behavioural findings are paralleled by facial electromyography (EMG) recordings from the zygomaticus muscle, which showed larger activity during fluent processing compared to dysfluent processing and may reflect the positive affect elicited by such processing. Response fluency also can elicit affect and influence judgements of stimuli. Affordance effects refer to the finding that responses to objects are facilitated when these responses are compatible with the actions used to manipulate those objects. So, for example, responses to an image of a cup are faster when the handle of the cup is oriented in the same direction as the response hand (Tucker & Ellis, 1998). This is known as stimulus-response (S-R) compatibility and it also elicits more positive affect than incompatibility as measured by facial EMG (Cannon, Hayes & Tipper, 2010). The elicitation of positive affect by compatibility can become associated with stimuli presented in close succession. Chinese ideographs were rated more positively after S-R compatible compared to incompatible trials (Regenberg, Hafner & Semin, 2012).

According to Rolls (2008), the uniqueness and variety in emotional experiences arises through six factors: (1) the reinforcement contingency, i.e. whether the stimulus is a reward or punisher or whether the emotion is elicited by its encounter, omission or termination. (2) The intensity of the reinforcer. (3) Multiple reinforcement associations. (4) Secondary reinforcers associated with different primary reinforcers that elicit the same type of emotion are experienced as qualitatively different. This would be due to the type of learning in which emotional reactions to CSs are mediated by a representation of the US. (5) Different secondary reinforcers may form associations with particular USs more or less easily and qualitative differences between them results in differences in the emotions experienced. (6) Behavioural responses that are available. For example, if an active behavioural response can prevent the omission of a reward, then anger may be produced. In contrast, if no behaviour can be performed to prevent the loss of reward, sadness may ensue.

Emotional processes in humans have been studied using a range of different emotion evoking stimuli such as from the international affective picture system (Bradley & Lang, 2000) or from standardised facial expression stimuli databases (Tottenham et al., 2009; Young, Perrett, Calder, Sprengelmeyer & Ekman, 2002). In humans, it has been repeatedly observed that emotionally arousing stimuli are more vividly and accurately remembered than neutral stimuli (Bradley, Greenwald, Petry & Lang, 1992; LaBar & Cabeza, 2006; LaBar & Phelps, 1998; Todd, Talmi, Schmitz, Susskind & Anderson, 2012). One theory that has been proposed to account for this effect is referred to as the modulation model (McGaugh, 2004). According to this theory, emotionally arousing stimuli elicit the secretion of stress hormones such as adrenaline and cortisol that activates the amygdala which in turn enhances the consolidation of the memory stored in other regions such as the hippocampus (Roozendaal & McGaugh, 2011). Whilst the memory modulation model provides a good account of long-term memory consolidation, particularly single trial conditioning studies in rodents, it is limited by several factors. First, consolidation takes hours. In contrast, most experiments on emotional memory in humans have used memory tests that are presented within minutes after learning. This is referred to as early long-term memory. Secondly, the emotional and neutral stimuli presented in memory experiments are done so in rapid succession, meaning, due to the long duration of stress hormone release during consolidation, such hormones should be active during both emotional and neutral trials and therefore no differences should be observed. These limitations of modulation theory have led to proposals of alternative accounts. Talmi (2013) has proposed that increased recruitment of attention, distinctiveness between emotional and neutral stimuli and increased semantic relatedness of emotional stimuli, which allows stimuli to be integrated into memory more easily, are more likely reasons for enhanced emotional memory effects. In support, Talmi and McGarry (2012) found that detection of auditory stimuli was reduced when presented concurrently on emotional compared to neutral encoding trials and that this accounted for the enhanced memory for emotional stimuli in free recall. Furthermore, instructions to pay more attention to neutral stimuli abolished the emotional memory effect, as did presenting the emotional and neutral stimuli in separate blocks, to reduce distinctiveness, whilst equating the neutral and emotional stimuli for extent of semantic relatedness between items.

There are large individual differences in aspects of emotion that underpin variability in temperament, personality and susceptibility to psychopathology. Different people can react to the same emotional stimulus in markedly different ways. Davidson (1998) has proposed several ways in which emotions may differ between individuals. Firstly, there are differences in the regulation of emotions. This includes ways in which individuals may voluntarily or automatically direct their attention or thoughts in a way that amplifies or diminishes the processing of the emotional stimulus. Secondly, individuals differ in their threshold for emotion elicitation. That is, the strength of the emotional stimulus needed to elicit the emotion and to elicit comparable levels of the emotion between different individuals will vary. There may also be differences in dose-response functions, where the change in strength of the elicitor needed to produce a step change in emotion will also vary. Thirdly, once an emotion has been elicited, individuals vary in the peak or amplitude of this response. Fourthly, individuals vary in their rise time to peak. Some individuals react more quickly than others. Fifthly, individuals vary as to how quickly it takes for the emotion to dissipate, referred to as recovery time. Finally, individuals may vary in their experience of different emotions and the channels through which these are expressed. For example, someone may be more likely to experience fear than happiness and show large changes in physiological activity that is not paralleled by equal changes in the subjective report of the emotion. Both rise time to peak and recovery time are aspects of affective chronometry, the timing of emotional responses. The role of timing in individual differences in emotion underscores the importance of methods that have high temporal resolution, such as EEG, to distinguish where emotional processing begins, peaks and ends. Where only the onset and offset of emotion differs between subjects such information could be missed by methods with poorer resolution such as fMRI where activity is often computed as the aggregate in long trials or blocks.

## Neural Correlates of Emotion

While primary and secondary reinforcing stimuli can be from several different sensory modalities, here I focus primarily on visual stimuli as this is the modality through which a great deal of human social cues are transmitted and which will be used in the investigations that form the basis of this thesis. Such stimuli are processed to a high level of complexity throughout the visual hierarchy from early visual cortex along the ventral visual pathways. The outcome of such processing is a representation that is invariant to changes in the appearance of the object due to changes in translation, position, size and luminance etc. The objects need to be represented at a high capacity, meaning that different objects need to be highly differentiable amongst each other, for example, to differentiate one face from another so that an appropriate response is consistently executed. At the same time, the emotional response to one stimulus needs to generalise to different stimuli when these stimuli share similar properties that lead to similar outcomes. For example, an angry facial expression pattern is similar across different individuals and usually implies a similar outcome, threat. As previously discussed in relation to one particularly important type of visual object, faces, this processing appears to take place in the visual areas of the occipital and temporal lobes such as the superior temporal sulcus, fusiform gyrus and anterior temporal cortex. Although these regions carry crucial inputs to emotion regions, they do not represent the value of the stimulus itself. Instead, inputs about the identity of a stimulus in visual cortices trigger a cascade of processing in orbitofrontal cortex, amygdala, ventral striatum, insula and anterior cingulate cortex that underpins an integrated emotional response. However, a plethora of evidence suggests that visual cortices are modulated by the intensity of the emotion through back projections from emotion regions that serve to increase attention and perception to stimuli of emotional significance (Vuilleumier & Driver, 2007).

## Orbitofrontal Cortex

The orbitofrontal cortex (OFC) represents the specific identity of outcomes and their subjective value, given the individuals current state (Rudebeck & Murray, 2014). Bilateral lesions to the OFC in humans can produce changes in subjective emotional state and social behaviour including affective blunting, impulsiveness and irresponsibility (Hornak et al., 2003). These findings suggest that orbitofrontal damage produces a deficit in the ability to respond with appropriate emotions and to use those emotions to learn which behaviours are appropriate in a given situation. FMRI in humans suggests the OFC represents the reward and punishment value of stimuli commonly encountered by humans. The medial OFC is activated by monetary reward compared to monetary loss (O'Doherty, Kringelbach, Rolls, Hornak & Andrews, 2001). This activity scales with the magnitude of the money gained/ lost on individual trials. In contrast, activity in a lateral area of the OFC is larger for monetary losses suggesting this area represents punishing stimuli. The OFC appears also to encode the value of social stimuli. Different neurons in the OFC in monkeys have been found to be selective for facial identity, facial expression and head orientation (Rolls, Critchley, Browning & Inoue, 2006). The activity of some of these OFC neurons scale with how much the monkey values the social signals observed as measured by the amount of juice reward the monkey is willing to forgo to see the image, independent of whether a choice is made (Watson & Platt, 2012). This representation of social value is not a general representation of reward as such neurons are independent to those that represent juice value. This activity also correlates with how much the monkey looks at the images suggesting they also index interest.

In humans, activity in the medial OFC is larger for attractive than unattractive faces (O'Doherty, Winston, Critchley, Perrett, Burt & Dolan, 2003). Like monetary value, this activity is positively correlated with individual ratings of the attractiveness of particular faces. A negative correlation with attractiveness ratings was also observed in the lateral OFC. The area of the medial OFC that responds to monetary and social reinforcement is overlapping. Lin, Adolphs and Rangel (2012) found activity in medial OFC was larger at the time when participants chose stimuli that predicted monetary or social reward in the form of a happy facial expression and also when the reward was received. The structure of OFC predicts the influence that other people can have on participant's evaluative judgements of stimuli. Campbell-Meiklejohn et al. (2012) found that grey matter volume in the lateral OFC predicted the participant's change in desirability ratings for pieces of music that were observed to be chosen or not chosen by a credible other person.

### Amygdala

The amygdalae comprise bilateral clusters of 13 interconnected nuclei located in the anterior medial temporal lobes. Early lesion studies showed bilateral excision of the amygdalae in monkeys produced tameness, emotional unresponsiveness including reduced aversive and aggressive behaviours, abnormal food preferences and a loss of social standing (Aggleton & Young, 2000; Weiskrantz, 1956). The amygdala is most commonly associated with processing aversive, punishing stimuli as it is has been strongly implicated in fear and fear conditioning (Ledoux, 2000). Inputs from all major cortical and subcortical sensory pathways converge in the lateral nucleus where cells representing conditioned stimuli (CS) synapse onto cells representing unconditioned stimuli (US). When the CS is experienced alone, it reactivates cells associated with the US, and, via connections with the central nucleus, triggers a range of fear responses in other brain regions such as the hypothalamus, which controls heart rate, the periaqueductal gray in the midbrain, which controls freezing, the stria terminalis which controls neuroendocrine responses, and the cholinergic basal forebrain nucleus of meynert, which controls cortical arousal (Ledoux, 2000). The role of the amygdala in processing aversive stimuli has been confirmed in humans. Patients with bilateral lesions to the amygdala show impairments in the experience, learning, expression, recognition and understanding of fear (Feinstein, Adolphs, Damasio & Tranel, 2011). For example, one such patient shows no fear of snakes, spiders and haunted houses, self-reports abnormally low levels of fear and anxiety (Feinstein et al., 2011), is impaired in fear conditioning (Bechara et al., 1995) and recognising fearful facial expressions expressed by others (Adolphs, Tranel, Damasio & Damasio, 1994). In studies of normal humans, the amygdala is the most commonly activated area in response to emotional faces and scenes (Sabatinelli et al., 2011) and has been the region most strongly implicated in mediating the influence of emotion on visual perception and attention. This is because it is commonly co-activated and correlated with parietal and visual cortex activity during viewing of emotional scenes or emotional facial expressions (Bradley et al., 2003; Vuilleumier & Pourtois, 2007). This is likely due to the large connections projecting from the amygdala to all visual regions in the occipital and temporal cortices (Freese & Amaral, 2005; 2009). As mentioned previously, this may also help facilitate the encoding of information into memory.

The amygdala’s central role in emotion means it also plays an important role in social cognition by decoding the valence of others’ social cues and orchestrating an appropriate emotional response. For example, there are neurons in the monkey amygdala that respond to facial identity, facial expressions and gaze (Gothard, Bataglia, Erickson, Spitler & Amaral, 2007; Leonard, Rolls, Wilson & Baylis, 1985; Mosher, Zimmerman & Gothard, 2014). Similar to the studies suggesting that the amygdala is particularly important for fear, studies of the social behaviour of amygdala lesioned monkeys have shown that these animals tend to have less nervous and fearful personality qualities, are less socially inhibited and show more affiliative behaviour than non-amygdala damaged control animals (Machado et al., 2008). This is consistent with the notion that the amygdala is involved in evaluating potential social threats or aversive social stimuli (Amaral, 2003). In humans, support for this notion has come from studies showing that amygdala lesioned patients rate unfamiliar faces as more trustworthy and approachable than normal individuals (Adolphs, Tranel & Damasio, 1998). In normal participants, amygdala activity also correlates with trustworthiness ratings of faces (Engell, Haxby & Todorov, 2007; Winston, Strange, O'Doherty & Dolan, 2002).

Although the amygdala is crucial for fear and aversive processing, it is important to bear in mind that more recent studies have demonstrated that cells code both the positive and negative value of primary reinforcers (such as juice reward and an aversive air puff to the eye) as well as associated secondary reinforcers (Baxter & Murray, 2002; Belova, Paton & Salzman, 2008; Paton, Belova, Morrison & Salzman, 2006).

### Ventral Striatum

The ventral striatum, particularly the nucleus accumbens, appears to be important for motivation and may be a region where the value of stimuli computed in the amygdala and OFC can influence behaviour. In monkeys, the activity of dopamine cells in the striatum appears to encode reward prediction and a reward prediction error learning signal during Pavlovian conditioning. Activity in response to a reward is greater when the reward is unpredicted (Schultz, 1998). As a conditioned stimulus is learned to predict the reward, the activity of the dopamine cell shifts to the presentation of the conditioned stimulus instead of the reward to predict its imminent contingent presentation. Similar activity related to anticipating rewards is observed in the ventral striatum of humans when anticipating monetary gain (Knutson & Greer, 2008). The vast majority of research investigating the ventral striatum has focused on reward learning. However, there is a growing literature showing that activity in the ventral striatum is also implicated in aversive learning and motivation (Seymour et al., 2004; Levita et al., 2009; Levita, Hoskin & Champi, 2012). Instead of coding the reward or punishment value of stimuli, the ventral striatum, particularly dopamine cells, have been proposed to represent "incentive salience", or, how much the animal "wants" the stimulus (Berridge & Robinson, 1998). The corresponding state of "aversive salience" is elicited by aversive stimuli and represents "wanting to avoid". This activity may help to form associations between behavioural responses and motivationally important stimuli, reduce the threshold for such responses and prepare these responses before the stimulus occurs by predicting their occurrence. Ventral striatal activity also appears to be important for motivation for social stimuli. Activity in this region is greater when participants anticipate the outcome of another person responding positively or negatively towards them before making a response to try and gain or avoid these outcomes, respectively (Kohls et al., 2013). In addition, response times in anticipation of positive social cues were correlated with the ventral striatal activity. In contrast, receipt of the outcome resulted in activation in the OFC and insula.

### Insula

The dorsal posterior insula receives inputs from the autonomic nervous system about the physiological state of the body such as the mechanical, thermal, chemical, metabolic and hormonal status of skin, muscle, joints, teeth and viscera (Craig, 2003). This interoceptive representation appears to be important for homeostatic emotions that reflect the survival needs of the body such as pain, thermoregulation, hunger, thirst etc. The anterior insula is activated in a wide range of circumstances, most pertinently of which is emotional awareness including feelings of love, anger, fear, sadness, happiness, disgust, aversion, unfairness, inequity, indignation, trust, empathy and social exclusion among others (Craig, 2009). It has been proposed that the mid and anterior parts of the insula re-represent the sense of the physiological condition of the body in the posterior insula whilst integrating with it information about the sensory environment, motivational, hedonic and social conditions represented in temporal cortex, amygdala, OFC, ventral striatum, anterior cingulate cortex, and prefrontal cortices (Craig, 2009). The insula would therefore appear to be an area in which emotion-related changes in physiological activity, that is somatic markers, are sensed and used to inform emotional states (Damasio, 1996).

### Anterior Cingulate Cortex

The anterior cingulate cortex (ACC) is often co-activated with the insula but is not necessarily implicated in awareness (Craig, 2009). It has been shown to respond to rewards and punishers (Rolls, 2008). It is implicated in the affective qualities of pain. For example, neurons respond to painful stimuli regardless of where on the body the pain originates from (Vogt, 2005). Lesions to this region reduce feelings of pain and avoidance (Ballantine, Cassidy, Flanagan & Marino, 1967). The ACC is also implicated in autonomic control by way of its connections to brain stem areas (Vogt, 2005; Nagai, Critchley, Featherstone, Trimble & Dolan, 2004), particularly in effortful and evaluative tasks (Critchley, 2009).

### EEG/ ERP Correlates of Emotion

Emotional modulations of ERPs have been observed in different time-windows and stages of stimulus processing and the effects appear to be more reliable among some components than others. Firstly, a larger centroparietal negativity has been observed during anticipation of an upcoming stimulus when a cue has been presented informing the participant that the stimulus will be positively or negatively valenced as opposed to neutral (Poli, Sarlo, Bortoletto, Buodo & Palomba, 2007). This is known as the stimulus-preceding negativity. After stimulus onset, some evidence suggests that there may be modulation of early visual ERP components such as the P1, N1, P2 and N2/ early posterior negativity (Hajcak, Weinberg, McNamara & Foti, 2012). In response to faces, evidence for modulation by emotional expression at the level of the N170 and vertex positive potential (VPP) is unclear. There appears to be variation across studies in the emotional expressions that elicit a difference from neutral (Hajcak et al., 2012) and some studies fail to find differences (Eimer & Holmes, 2007; Eimer, Holmes & McGlone, 2003; Holmes, Winston & Eimer, 2005). While the detection of significant effects could reflect emotional modulation of perception, mediated by the amygdala, there is also the possibility that the significant differences observed are simply due to differences in the perceptual properties between the emotional and neutral stimuli given the role of these components in early visual processing.

In contrast to earlier visual components, the P300 and late positive potential (LPP) have been shown more reliably to be sensitive to stimuli with positive and negative valence compared to neutral stimuli. The P300 is typically measured as a positivity on parietal sites between 300 and 500 milliseconds after stimulus onset. The P300 has classically been studied using the oddball paradigm where the ERP signal is larger to infrequent target stimuli compared to frequently presented standard stimuli (Polich, 2012). This suggests that the P300 reflects the deployment of attentional resources to the processing of salient stimuli in the environment such as when a change in an ongoing stimulus occurs. As the P300 is also larger to emotive compared to neutral stimuli (Keil et al., 2002), this suggests that these stimuli automatically garner attention and act as natural targets (Hajcak, McNamara & Olvet, 2010).

The LPP is a slow wave beginning at the level of the P300 with a parietal maximum, which is larger for positively and negatively valenced stimuli compared to neutral, and usually remains sustained for the duration of the stimulus and even after stimulus offset (Cuthbert, Schupp, Bradley, Birbaumer & Lang, 2000; Keil et al., 2002; Palomba, Angrilli & Mini, 1997; Schupp et al., 2000). In EEG source localisation and combined EEG-fMRI studies, the LPP has been shown to relate to concurrent activity in brain regions involved in visual/ attentional processing such as lateral occipital, parietal and inferotemporal cortices and emotion regions such as the OFC, insula, anterior cingulate cortex, ventral striatum and amygdala (Harris, Adolphs, Camerer & Rangel, 2011; Liu, Huang, McGinnis-Deweese, Keil & Ding, 2012; Moratti, Saugar & Strange, 2011; Sabatinelli, Keil, Frank & Lang, 2013; Sabatinelli, Lang, Keil & Bradley, 2007). The LPP is believed to reflect emotion processing and the representation of value, which guides increased sustained attention to, and memory encoding of, the affective stimulus (Hajcak et al., 2010). While a concurrent attention demanding working memory task reduces the LPP in general, a larger LPP for affective compared to neutral pictures is unaffected by load (MacNamara, Ferri & Hajcak, 2011). Emotion LPP effects are also unaffected when a difficult math task is performed concurrently compared to an easy math task or no task at all (Hajcak, Dunning & Foti, 2007). This suggests that the LPP reflects the way in which emotional stimuli command attention. The only way in which cognitive processes appear to modulate the LPP is when they involve conscious attempts to regulate emotion. The LPP has been shown to be reduced when attention is cued to a non-arousing portion of unpleasant pictures (Dunning & Hajcak, 2009; Hajcak, Dunning & Foti, 2009). The LPP to unpleasant images is also reduced after cognitive reappraisal, where the image is reinterpreted in a less negative way (Foti & Hajcak, 2008; Hajcak & Nieuwenhuis, 2006; Moran, Jendrusina & Moser, 2013).

The bivalent way in which emotion modulates the LPP is similar to the way in which memory for such stimuli is enhanced (Dolcos & Cabeza, 2002) and an LPP is elicited when viewing stimuli correctly recognised as previously shown in the presence of emotional compared to neutral stimuli (Smith, Dolan & Rugg, 2004), suggesting that the emotional qualities of the stimulus are retrieved and act as retrieval cues along with other properties of the stimulus (Buchanan, 2007). Although the LPP is modulated by both positively and negatively valenced stimuli, there is some evidence that negatively valenced stimuli may weigh more heavily on the brain, particularly when the stimuli are unexpected, as in oddball paradigms the LPP tends to be larger for negative stimuli that are comparable in magnitude to positive stimuli in terms of the absolute value of the difference in ratings from neutral (Ito, Larsen, Smith & Cacioppo, 1998; Hilgard, Weinberg, Proudfit & Bartholow, 2014; Wood & Kisley, 2006). The reason for this is proposed to be similar to the reason why the amygdala is more strongly responsive to threat-related stimuli. The cost of not orchestrating an appropriate response to negative stimuli is more likely to have greater negative consequences than not responding to positive stimuli of equivalent magnitude.

The LPP also appears to encode emotional responses to social cues. The LPP is larger for images containing faces compared to non-faces (Ferri, Weinberg & Hajcak, 2012;

Weinberg & Hajcak, 2010) and is sensitive to facial expressions of fear and anger (Smith, Weinberg, Moran & Hajcak, 2013), trustworthiness (Marzi, Righi, Ottonello, Cincotta & Pia Viggiano, 2012; Yang, Qi, Ding & Song, 2011) and attractiveness (Wiese, Altmann & Schweinberger, 2014). The LPP is believed to reflect emotional arousal as opposed to valence as it is correlated with arousal ratings (Cuthbert et al., 2000). However, given that studies have only looked at linear correlations, it could also show a quadratic trend with valence ratings. Thus, it most probably reflects both valence and arousal. Unlike fMRI and cellular recordings, the EEG reflects the aggregate of activity in the whole brain; therefore it is unsurprising that signals for both negative and positive affect are picked up.

Nevertheless, some coarse laterality distinctions have been made. Cunningham, Espinet, DeYoung and Zelazo (2005) found that words which denoted good socially relevant concepts elicited a larger LPP on left frontal electrodes compared to bad concepts. In contrast, on right frontal electrodes, the LPP was larger to bad compared to good judgements. These effects were present when the participants' task was to make non-evaluative (abstract vs. concrete) judgements of the stimuli, but was larger when they were making evaluative judgements. These findings are consistent with a well established strand of research which has established that greater left compared to right hemisphere alpha band (8-12 Hz) power desynchronisation is related to approach related motivation whereas the reverse is associated with avoidance related motivation (Harmon-Jones, Gable & Peterson, 2010). Decreases in alpha power are associated with increased metabolic brain activity as measured with positron emission tomography (PET; Cook, O'Hara, Uijdehaage, Mandelkern & Leuchter, 1998). Therefore these findings are consistent with the notion that left hemisphere frontal activity is related to positive affect and right hemisphere frontal activity is related to negative affect. The majority of these studies have been concerned with individual differences, showing that greater resting state left frontal activity is associated with trait positive affect and increased positive affect to positive stimuli whereas the inverse is associated with trait negative affect and larger negative affect to negative stimuli (Tomarken, Davidson, Wheeler & Doss, 1992; Wheeler, Davidson & Tomarken, 1993). However, regardless of personality, lesions to the right frontal cortex have been shown to produce manic symptoms whereas left frontal lesions can result in depression (Black, 1975; Gainotti, 1972; Gasparrini, Satz, Heilman & Coolidge, 1978; Robinson & Price, 1982; Sackeim et al., 1982). In addition, injection of barbiturates into the left or right carotid arteries, which inhibits activity selectively in the ipsilateral hemisphere, has been found to induce depression and euphoria, respectively (Trimble, 2010). Interestingly, the performance of a task with the right hand, leads to greater positive affect and this correlates with greater left frontal activity (Harmon-Jones, 2006). EEG activity measured in terms of the LPP and alpha band desynchronisation from parietal electrodes are highly correlated in response to affective pictures (De Cesarei & Codispoti, 2011).

## Psychophysiological Correlates of Emotion

Emotional stimuli evoke changes in facial expression, pupil size, skin conductance and heart rate as measured with electromyography (EMG), pupilometry, galvanic skin response (GSR) and the electrocardiogram (ECG). Pupil size and skin conductance are correlated with emotional arousal, being comparable in amplitude to both positive and negatively valenced stimuli (Bradley, Miccoli, Escrig & Lang, 2008; Lang, Greenwald, Bradley & Hamm, 1993). In contrast, facial EMG recorded from the corrugator muscle can index the valence of the emotional stimulus. The contraction of the corrugator is a crucial action unit for the expression of anger and distress (Ekman & Friesen, 1978). Strong evidence for the relation of corrugator EMG activity to negative emotion is demonstrated by the reliable correlations with ratings of affect in response to affective pictures. Corrugator activity increases linearly as ratings become more negative and decreases as ratings become more positive (Greenwald, Cook & Lang, 1989; Lang et al., 1993; Larsen, Norris & Cacioppo, 2003), making it the most useful measure of the valence of emotional reactions. It is also correlated with amygdala responses to affective pictures, which, as mentioned, is commonly activated by negatively valenced stimuli (Heller, Greischar, Honor, Anderle & Davidson, 2011; Heller, Lapate, Mayer & Davidson, 2014).

In contrast, the zygomaticus major muscle shows a bivalent response profile, correlating with affective ratings of both positive and negative stimuli (Greenwald et al., 1989; Lang et al., 1993; Larsen et al., 2003) and is involved in both smiling and grimacing (Bradley & Lang, 2000; Ekman & Friesen, 1978). Another facial muscle that appears to be particularly active when expressing negative emotions is the levator labii muscle, which is active during expressions of disgust (Ekman & Friesen, 1978). EMG activity recorded from the levator is greater when participants experience disgust induced by drinking an unpleasant tasting solution, observing disgusting images or when a partner in an economic exchange offers an unfair split of money to the participant (Chapman, Kim, Susskind & Anderson, 2009). In all cases, EMG activity correlates with self-reported disgust. These findings are supported by a further study, which demonstrated that corrugator and levator activity reflect the negative emotions elicited by different types of moral violations. Corrugator activity was found to be larger when participants considered third person statements that described behaviours that violated moral norms regarding harm. In contrast, levator activity was larger for behaviours that violated notions of purity and fairness (Cannon, Schnall & White, 2011). This study suggests the corrugator is also activated by negatively valenced social stimuli.

It is a well established and reliable finding that observation of facial expressions of anger can evoke activity in the corrugator of the observer whereas observation of happiness triggers activity of the zygomaticus (Dimberg, 1982; Dimberg & Lundqvist, 1990; Dimberg & Petterson, 2000; Dimberg & Thunberg, 1998; Dimberg, Thunberg & Elmehed, 2000; Dimberg, Thunberg & Grunedal, 2002). Similar findings were obtained with bodily and vocal expressions (Magnee, Stekelburg, Kemner & de Gelder, 2007; Tamietto et al., 2009). Some researchers have proposed that the similarity of the expression between the observer and expresser is evidence of mimicry and that mirror neurons are involved in understanding others' emotions (Niedenthal, 2007). However, there is also the possibility that the congruent EMG reactions may be due to an emotional reaction to the facial expressions of others, similar to that observed with affective pictures, as opposed to motor mimicry per se. Evidence supports the latter of these theories. For example, when the faces to be mimicked have been associated with monetary loss or unfairness in an economic transaction there is larger activity of the corrugator and zygomaticus when the face expresses anger than when the person has behaved fairly or is associated with monetary gain (Hofman, Bos, Schutter & van Honk, 2012; Sims, Van Reekum, Johnstone & Chakrabarti, 2012). Likewise, zygomaticus activity in response to happy faces is larger when associated with monetary gains (Sims et al., 2012). In addition, when stimuli are presented very rapidly (~17 ms), below conscious awareness, happy faces still elicit larger activity of the zygomaticus and angry faces still elicit larger activity of the corrugator despite the fact that participants have no information about the detailed pattern of facial muscle activity in the observed face (Neumann, Schulz, Lozo & Alpers, 2014). Furthermore, surprised faces, which can be interpreted as either positive or negative in valence, have been shown to elicit larger corrugator activity among individuals who judge them to be negative compared to those who judge them to be positive (Neta, Norris & Whalen, 2009). Mimicry-like EMG activity has also been observed in response to neutral faces whose identity has been contingent upon angry or happy facial expressions in previous trials, but only in participants who showed EMG mimicry patterns to these expressions (Aguado et al., 2013). The neutral faces that had expressed anger compared to happiness were also rated more negatively. Corrugator activity in response to moderately pleasant pictures is reduced when presented amongst mildly pleasant pictures compared to when presented among extremely pleasant pictures (Larsen &

Norris, 2009). This suggests that the evaluation and emotional reaction to one stimulus can be biased by the intensity of other stimuli presented in close proximity. Therefore there may be a phenomenon of affective contrast or comparison processes in evaluative cognition whereby the emotional response to one stimulus is in reference to other stimuli, similar to influential companions effects observed in behavioural studies (Poulton, 1982). In sum, these studies show that valenced social cues can trigger emotional facial reactions, which may be expressive or communicative in nature and could index the learned reputation of the observed face. This may be greater if the range of face valences presented in a paradigm is large.

It is important to note that different psychophysiological measures are correlated and reflect underlying dimensions of emotional experience. Factor analyses of data from multiple behavioural and psychophysiological measures in response to affective pictures have shown that corrugator EMG, heart rate and valence ratings load most strongly on to one of the two factors that account for the most variance and that this is separate to a second factor upon which LPP, skin conductance and arousal ratings load onto (Cuthbert et al., 2000).

## Trustworthiness

Trustworthiness can be operationally defined as “a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behaviour of another” (Rousseau, Sitkin, Burt & Camerer, 1998). Trustworthiness judgements of other people have been studied in one of two different ways. The first way in which trust has been studied is in terms of learning. Trustworthiness is manipulated by varying the association between particular individuals and their behaviour, typically in an economic exchange, or via biographical information that is believed to describe the individual or other social cues such as facial expression or gaze. The most common paradigm used for this purpose is the trust game or a variation thereof. In this game, one individual, the proposer, is given an amount of money (say £10) and is asked to choose how much they would like to invest this money in a trustee. This investment is then tripled and the trustee can then choose to split the money in either a fair or unfair way (Fehr & Camerer, 2007). In this way both the propensity to trust and potential to behave in a trustworthy manner can be investigated when the participant is the proposer or the trustee, respectively.

The second way in which trustworthiness has been studied is by simply examining trustworthiness ratings of faces that vary widely in appearance. Oosterhof and Todorov (2008) acquired ratings of computer-generated faces on 15 of the most commonly used adjectives used to describe faces, including trustworthiness. The ratings were then submitted to a principal components analysis (PCA) to examine whether there are similarities between the ratings using different adjectives that might be explained by a smaller set of underlying dimensions. It was found that two principal components accounted for the majority of the different ratings. The first component, which explained most of the variance (~63%), was defined as valence and it was most strongly correlated with trustworthiness. Subsequently, computer generated faces were manipulated so that their features correlated with ratings of trustworthiness ranging from extremely trustworthy to extremely untrustworthy. It was found that extremely untrustworthy faces were judged more masculine and angry whereas extremely trustworthy faces were judged more feminine and happy. This has led to the proposal that judgements of the trustworthiness of neutral faces are based on an overgeneralisation of systems tuned to evaluate emotional facial expressions via similarity in appearance (Engell, Todorov & Haxby, 2010; Said, Sebe & Todorov, 2009). Despite large agreement both within and across raters in trust judgements of faces, evidence for the validity of these judgments appears to be inconsistent (Rhodes, Morley & Simmons, 2012; Rule, Krendl, Ivcevic & Ambady, 2013; Stirrat & Perrett, 2010).

Trusting another person has the potential to lead to rewarding or punishing outcomes or the omission of such outcomes depending on whether the person is trustworthy or untrustworthy. As such, it is not surprising that a strong role for emotion has been implicated in trust learning, judgements and decisions. Inducing participants to feel happy leads to increased ratings of trustworthiness of unfamiliar individuals compared to inductions of anger (Dunn and Schweitzer, 2005). Many of the facial EMG, SCR, ERP and fMRI correlates of emotion show similar responses to the trustworthiness of other individuals as they do various other emotional stimuli. The proposal of unfair offers by a trustee in a trust game elicited larger EMG activity of the levator muscle (Chapman et al., 2009) and skin conductance responses (van't Wout, Kahn, Sanfey & Aleman, 2006) than fair offers and corrugator activity was larger to faces who had consistently behaved unfairly in such interactions (Hofman et al., 2012). Similar findings have been observed with ERPs. Specifically, a feedback negativity at 270-370 ms, which has a maximum on central electrodes, is larger when participants receive an unfair compared to a fair split of money (Osinsky, Mussel, Ohrlein & Hewig, 2013). Previous studies have suggested that this reflects a good/ bad evaluation of the outcome as an error (Miltner, Braun & Coles, 1997). After three blocks of learning, a larger negativity between 270-370 ms was also observed during initial presentation of unfair faces before their proposal for the split of money was presented. Using fMRI, it has been shown that unfair offers in trust games lead to larger activity in the ACC and insula; activity in this latter region correlates with acceptance rates of these offers (Sanfey, Rilling, Aronson, Nystrom & Cohen, 2003). Similar to studies of reward prediction error learning with non-social stimuli, the striatum has been implicated in the learning of trust in the trust game (King-Casas et al., 2005). Early in the experiment, larger striatal activity of trustee participants was observed in response to the proposals of a trustworthy compared to an untrustworthy proposer. However, later in the experiment, as trust was learned, this difference in activity shifted to when the trustee was anticipating the proposers decision.

The trustworthiness of an individual can fluctuate across multiple interactions. Behrens, Hunt, Woolrich and Rushworth (2008) found that prediction errors in trustworthiness, defined as a change in the validity of advice given by a confederate about a choice that led to monetary gain or loss, elicited larger activity in the pSTS, dorsomedial prefrontal cortex and temporoparietal junction (TPJ). These regions have previously been implicated in representing others' mental states and understanding others intentions (den Ouden, Frith, Frith & Blakemore, 2005; Saxe & Kanwisher, 2003; Saxe, Xiao, Kovacs, Perrett & Kanwisher, 2004). The dorsomedial prefrontal cortex has also been consistently implicated in the learning of trust from behavioural descriptions (Baron, Gobbini, Engell & Todorov, 2011; Mende-Siedlecki, Cai & Todorov, 2013). Behrens et al. (2008) also found that ACC tracked the changeability of the confederate's trustworthiness and that this signal was correlated with the influence of the advice on choices and reward anticipation activity in the vmPFC. Although not activated in trust games using fMRI, the amygdala may also play a role in learning trust. Patients with amygdala lesions show a pattern of benevolent reciprocity whereby they give more money to trustees who behave unfairly than normal participants (Koscik & Tranel, 2011; van Honk, Eisenegger, Terburg, Stein & Morgan, 2013).

Neural correlates of emotion have also been implicated in judgements of trustworthiness from the appearance of the face alone. The amygdala appears to be the key region implicated in the coding of trustworthiness from facial appearance. Patients with amygdala damage have been shown to judge faces as trustworthy, which neurologically normal participants judge as untrustworthy (Adolphs et al., 1998). However, results across fMRI studies are somewhat inconsistent as some studies suggest that the amygdala codes only untrustworthiness whereas other studies suggest it codes both trustworthiness and untrustworthiness. The amygdala response has been shown to be larger for faces that are rated untrustworthy compared to trustworthy and amygdala activity has been shown to increase linearly as trustworthiness ratings of the faces decrease (Engell et al., 2007; Winston et al., 2002). However, other studies found a quadratic or U-shaped amygdala response that increased to both trustworthy and untrustworthy faces relative to neutral faces (Mattavelli, Andrews, Asghar, Towler & Young, 2012; Todorov, Baron & Oosterhof, 2008). Both linear and quadratic effects have been observed when the faces are presented subliminally and this may be related to processing in different amygdala subregions or the use of blocked compared to event-related designs (Freeman, Stolier, Ingbretsen & Hehman, 2014).

A relation between emotion and trustworthiness evaluations has also been implicated by EEG/ ERP studies. Marzi, Righi, Otonello, Cincotta and Pia Viggiano (2014) used computer generated trustworthy, untrustworthy and neutral faces created by varying the face dimensions identified by Oosterhof and Todorov (2008). They found that the LPP was larger to faces classified by participants as untrustworthy compared to trustworthy. The LPP was also larger to untrustworthy faces when only considering faces whose classification as trustworthy or untrustworthy was congruent with the predefined trustworthiness category to which it belonged. These effects appeared to be specific for trustworthiness judgements. The same effects were not found when comparing faces, which the participants indicated they would vote for compared to those they would not vote for in a political election. Yang, Qi, Ding and Song (2011) used the same faces as Marzi et al. and also found that the LPP was larger to untrustworthy compared to trustworthy faces. In addition, using a regression analysis, Yang et al. found that larger LPP amplitudes predicted lower trustworthiness ratings. As discussed previously, the LPP has sometimes been found to be larger for negatively valenced stimuli compared to positively valenced stimuli which are comparable in terms of the difference in the absolute value of valence and arousal ratings from neutral stimuli (Hilgard et al., 2014; Ito et al., 1998; Wood & Kisley, 2006). As such, these findings lend further support to the existence of a negativity bias.

As previously described, memory is often enhanced for emotional stimuli. Given the above data suggesting a role for emotion in learning, judgements and decisions about trust, it is not surprising that trust can impact upon memory. Memory for untrustworthy faces is better than for trustworthy faces regardless of whether this is based on judgements of facial appearance or on information learned from positive or negative behavioural descriptions (Mattarozzi, Todorov & Codispoti, 2014; Rule, Slepian & Ambady, 2012). Learned mistrust is also more resistant to extinction than trust, when participants are instructed to disregard previously learned associations between faces and trust (Suzuki, Honma & Suga, 2013). These findings also suggest a negativity bias in memory perhaps by the known role of emotion in enhancing the encoding of stimuli via attention (Talmi, 2013).

## The Role of Emotion in Learning Trustworthiness from Gaze-Cues

The review of the previous research described provides a framework within which hypotheses can be formed regarding the mechanisms through which trust is learned from identity-contingent gaze-cues. In the previous section we have seen that trust judgements are affected by emotion and the process of forming judgements of trust relies upon emotion related neurocircuitry. The same processes through which trust is formed from unfair/ fair behaviour in the trust game or through association with behavioural descriptors could also be recruited during learning of trust from incidental task irrelevant gaze-cues. Indeed, several studies have shown that congruent and incongruent gaze-cues do activate emotion related neurocircuitry. Using fMRI, Schilbach et al. (2010) used a paradigm in which participants were instructed to either follow or not follow the gaze of another individual towards one of three objects, or were instructed to direct the other individual's gaze towards an object using their own gaze, which was either reciprocated or not reciprocated by the other individual based on eye tracking. Participants believed that a real person was controlling the gaze of the face observed on the monitor. Schilbach et al. found that self-initiated joint attention compared to non-joint attention was rated more pleasurable, less difficult and elicited greater reward/ motivation related ventral striatum activity, which correlated with subjective pleasantness ratings. The pupil size of participants was also larger on non-joint compared to joint attention trials. Gordon, Eilbott, Feldman, Pelphrey and Vander Wyk (2013), used a very similar design with eye tracking but only examined self-initiated joint attention with a real face instead of an animated face and silhouettes of another person as the target object. The results were consistent with Schilbach et al. In response to joint attention compared to non-joint attention, larger activity was observed in the striatum, ACC, parietal cortices and fusiform gyrus. The reverse contrast showed larger activity in the temporoparietal junction. Interestingly, the activity in the striatum was driven as much by a decrease to non-joint attention as it was an increase to joint attention, suggesting that joint attention may increase reward and non-joint attention may decrease reward. These findings are consistent with the notion that shared attention and intentionality -the awareness of perceiving the same object and having the same goal as another individual- may be inherently rewarding and motivate social interactions (Tomasello, Carpenter, Call, Behne & Moll, 2005).

However, while neuroimaging studies do show gaze-cues elicit emotion, as of yet, no studies have examined the role of emotion in the learning of trust from identity-contingent gaze-cues. The experiments conducted as part of this thesis aim to rectify this omission. It is hypothesised that implicit emotional reactions to gaze-cues mediate the learning of trustworthiness. Specifically, the non-joint attention, attentional dysfluency or perceived deceptiveness of the incongruent gaze-cues will lead to a more negative emotional reaction compared to congruent gaze-cues which may be experienced as more rewarding or less aversive due to joint attention, fluency or cooperation. Signals relating to such implicit emotional reactions may be recorded using electroencephalography (EEG) and facial electromyography (EMG), which may be able to distinguish the valence of emotional reactions to gaze-cues as trust is learned and may be related to subjective judgements of trustworthiness. Such measures may also index the retrieval of emotion states learned from gaze-cues when participants simply observe the faces without gaze-cues.

In addition to providing a measure of emotional reactions during identity-contingent gaze-cueing, the facial EMG signal may also be indicative of somatic markers that are causally related to learning of trust. That is, facial expressions are not just an epiphenomenon of experiencing an emotion. Instead, participants rely on their embodied states to infer their emotional response to the faces and this information is used to guide judgements of trust (Schwarz & Clore, 2003). This may be similar to the type of trait judgement learning observed with other actions. For example, when participants make a response with the foot to indicate that they recognise another person who is playing football, they respond faster than when making a response with the finger, whereas they are faster at responding with the finger compared to the foot when recognising a different person typing. After many such trials, participants later judge such people as more sporty and academic, respectively (Bach & Tipper, 2007). In the case of gaze-cueing, the predominant information that is stored, associated and use to guide judgements may be from emotion rather than motor states.

Another method through which a role for emotion may be determined is through the induction of emotion. As the current investigations pertain to emotional learning involving social stimuli, the induction of social emotions elicited by priming of inclusion or exclusion might be expected to be influential. It is known that feeling excluded from a social interaction can elicit negative emotions (Blackhart, Nelson, Knowles & Baumeister, 2009; Seidel et al., 2013). The negative feelings induced by exclusion would be expected to amplify the effect of incongruent gazes-cues on learning of mistrust compared to inclusion. It is also crucial to determine whether the mediating role of emotion in learning face evaluations from gaze-cues is restricted to trust judgements and social stimuli. Trust and gaze are highly social. It would be expected that the effects of gaze on emotion and trust are domain-specific to social processing and not the result of a general process whereby attentional cueing by any stimulus can elicit changes in judgements such as liking. Therefore it will be important to examine whether the effects obtained with gaze and trust also generalize to arrows and liking.

Each of these issues was examined in a series of six experiments. Experiments 1 and 2 address the core issue of whether emotion plays a mediating role in the learning of trustworthiness from identity-contingent gaze-cueing using EEG and facial EMG as continuous and implicit measures of emotional responses. Experiment 3 addresses the role of emotion in the learning of trust from gaze-cues from a different angle. This experiment investigated whether increasing a participants positive or negative emotional state through a mood induction procedure involving the priming of social inclusion or exclusion would make participants more sensitive to the learning of trust from gaze-cues and boost any EMG responses observed in experiment 2. There are many ways in which social rejection can be elicited (Williams, 2007). In order to strongly elicit social exclusion, we used three different manipulations to prime both third party and first-person exclusion. Firstly, to prime third-party exclusion, we used movies depicting interacting shapes, which are typically attributed human like mental states. Some of these shapes are perceived to reject another shape (Over & Carpenter, 2009). Secondly, we presented real images of humans that depicted individuals being excluded. Thirdly, first person exclusion was primed by asking participants to introspect upon a time when they felt excluded (Uskul & Over, 2014). These manipulations were compared with closely matched inclusion primes. Experiment 4 followed up the results of experiments 1 to 3 by investigating the influence of pre-established face appearance based judgments of trustworthiness on the learning of trust from congruent and incongruent gaze-cues to provide information about underlying learning mechanisms. Experiment 5 investigates whether the effects observed in previous experiments are specific to the social qualities of faces, gaze and trust or is due to domain-general attentional cueing by comparing the learning of liking evaluations of faces from gaze with arrows which cue attention in a similar way to gaze. Finally, experiment 6 examines whether the learning of liking from gaze relates in a similar way to emotion-related ERPs as when trust is used to evaluate the faces.

# Chapter 2

## Experiment 1: Electroencephalographic (EEG) evidence for the role of emotion in learning trustworthiness from eye-gaze cues

In experiment 1, we address the core question of whether the learning of trust from identity-contingent gaze-cues is mediated by emotion by recording high-density EEG during an identity-contingent gaze-cueing task. As of yet, identity-contingent gaze-cueing studies have not been able to decipher whether the differences in trust judgements in response to congruent and incongruent gaze-cues is due to a more positive change in trust to congruent faces, a more negative change in trust to incongruent faces or both (Bayliss et al., 2009; Bayliss & Tipper, 2006). Therefore, in our paradigm, we measure ratings both before and after cueing instead of just forced choice two-alternative trustworthiness decisions after cueing. This allows for the control of any initial variability in trust judgements between conditions pre-cueing and for the assessment of the direction of the effect. Because of our ambivalence with respect to whether the cueing effects on trust are mediated by the positive emotions elicited by congruent faces or the negative emotions elicited by incongruent faces, EEG is the most suitable technique as the late positive potential (LPP) is sensitive to stimuli with both positive and negative valence (Cuthbert et al., 2000; Schupp et al., 2000). If both types of gaze-cues mediate learning of trust, the LPP could cancel out. Nevertheless, one may predict asymmetrical anterior activity patterns as positive stimuli elicit larger LPPs on left than right anterior electrodes and negative stimuli vice versa (Cunningham et al., 2006). One may also predict a negativity bias, whereby incongruent faces may elicit a larger LPP compared to congruent faces due to the greater weight given to negative information (Ito et al., 1998; Hilgard et al., 2014). In addition, during cueing, the pupil size of the faces was manipulated to be larger for congruent faces and smaller for incongruent faces. Pupil size is related to emotional arousal and interest in visual stimuli (Bradley et al., 2008). Pupil size was manipulated as an attempt to enhance effects of validity on trust learning under the assumption that small pupils signal a lack of interest and larger pupils greater interest and these states may be recognised and integrated with gaze-cues.

Given its role in emotion processing, we hypothesised that the LPP would index the learning of trust judgements from gaze-cues. Unlike evoked sensory components, such as the N170, the LPP is much more variable in the time domain and differences between conditions can occur in brief time-windows from 300 ms after stimulus-onset until stimulus-offset

(Hajcak et al., 2010). Thus, traditional ERP analyses may risk overlooking important effects. This is especially the case since our paradigm is relatively novel to EEG and because of the long duration and multiple trial periods in our design. Therefore, we used the statistical parametric mapping (SPM) approach (Kilner & Friston, 2010), in which analysis is performed on interpolated 3D images of scalp activity over time and corrected for multiple comparisons with random-field theory (RFT) to identify clusters of significant activity localised in time on the scalp. This approach avoids the bias associated with traditional ERP analyses (Ibanez et al., 2012; Kilner, 2013; Kriegeskorte, Simmons, Bellgowan & Baker, 2009), preserves the high resolution and dimensionality of the data whilst allowing us to detect significant effects with no a priori prediction about where and when these effects would occur. Neural responses to the faces before and after the identity-contingent gaze-cueing phase were also examined to determine whether effects of learning trust from gaze-cues modulates face related ERP components such as the P1, N170 and N250, which have been implicated in the perceptual processing, structural coding and recognition of faces, respectively (Schweinberger, 2011). The N250, a negative deflection at 250 ms on occipitotemporal electrodes, which is greater when preceded by an image of the same face (Schweinberger et al., 2002; 2004) or when the face has become familiar by repeated viewing (Pierce et al., 2011; Tanaka et al., 2006; Zimmermann & Eimer, 2013), was of particular interest given its proposed role in face recognition and may link particular faces with particular traits. Thus, recording neural activity before gaze-cueing provides an initial baseline to compare with after cueing, when faces are familiar and trust is learned. To investigate if individual differences in emotion relate to the effects of validity on trust judgements, participants were divided into those who showed the trust effect (congruent > incongruent) and those who did not. This variable was entered into all analyses as a between-subjects factor/ covariate. We also followed up strong effects on the scalp with Bayesian 3D source reconstruction using Multiple Sparse Priors (Friston et al., 2008), which previous studies have used to estimate the source of value signals (Harris et al., 2011; Larsen & O'Doherty, 2014) and face-responsiveness (Henson, Mouchlianitis & Friston, 2009).

## Method

### Participants

There were 26 participants overall of which 24 were female and all right-handed. Participants were volunteers from Bangor University with an average age of 22 (*SD*=4). All were neurologically normal with normal or corrected-to-normal vision and received course credit or £15 for taking part. The university ethics board approved all procedures.

**Stimuli and Apparatus**

The 16 faces used in the experiment were 8 female and 8 male faces adapted from the NimStim face database (Tottenham et al., 2009). A previous study has shown effects of gaze-cues on learning trust is greatest when the faces express happiness as opposed to a neutral expression (Bayliss et al., 2009). However, EMG research has shown zygomaticus activity in response to happy faces (Cannon, Hayes & Tipper, 2009; Dimberg, Thunberg & Elmehed, 2000). Therefore, in our subsequent EMG experiments, a high intensity expression would confound the measurement of responses to gaze. In addition, mildly happy faces have typically been used as neutral baseline comparison stimuli in experiments of facial expression perception due to the tendency for people to smile slightly in normal social interactions, where faces that show no contraction of the facial muscles could appear hostile (Mattavelli et al., 2013). Therefore, all faces were made to appear to have a mildy happy facial expression. The mildly happy faces were created by morphing a neutral version of each face with a happy version to create 20 frames varying from neutral to happy. A set of ten observers were then asked to adjust each face to the point at which it could just be detected as happy. The average frame chosen was used in the experiment. It is noteworthy that a previous experiment using the same faces as the current experiment found that happy expressions did not elicit a larger LPP compared to neutral (Smith et al., 2013). Thus, it is unlikely that the happy expression of the face will produce a ceiling effect in the LPP that limits any responses related to gaze.

The faces were divided into two groups, A and B. The faces in each group were matched for visual appearance and ratings of trust and attractiveness in a previous study (Bayliss et al., 2009). Leftward and rightward gaze-cues were created by moving the irises into the left and right hand corners of the eyes. In addition, for the cueing phase only, the pupil size of congruent faces was increased to appear dilated whereas the pupils of incongruent faces were made to appear constricted. The target stimuli were a set of 32 garage and 32 kitchen objects. There were 16 unique objects in each category, which were in two different orientations. All were blue coloured and presented centrally to the left or right hand side of the face in line with the eyes. The faces were 300 × 385 pixels in size and the objects were 175 × 175 pixels in size. The cueing phase was presented at a resolution of 800 × 600 pixels whereas the viewing and rating phases were presented at 600 × 480 pixels. The experiment took place in a Faraday cage to shield external electromagnetic noise and was maintained at a slightly cool temperature to avoid sweat waves. The experiment was run using E-Prime 1.0 (Psychology Software Tools, Inc, Sharpsburg, Pennsylvania, USA) on a 24” Samsung SyncMaster BX2431 LED display, which was 569 × 342 mm in dimensions and had a 500 Hz refresh rate.

### EEG Recording

Electroencephalographic data was collected continuously using a 128 channel BIOSEMI ActiveTwo system at 2048 Hz. All participants washed their hair with baby shampoo before suitable sized electrode caps were fitted. Gel was injected into each of the receptor sites before attachment of Ag-AgCl active electrodes. Horizontal eye-movements were recorded with electro-ocular (EOG) electrodes placed on the outer canthi of both eyes and vertical eye-movements with two electrodes each placed infraorbitally and supraorbitally around the left eye. The EEG was monopolar referenced online using a common mode rejection active electrode.

## Design and Procedure

### Trust Rating Phases

Before the initial viewing phase and after the end viewing phase, participants completed the rating phases. As in the viewing phases, both initial and final rating phases were the same. Each trial began when participants pressed spacebar, at which point a fixation cross appeared for 1000 ms followed by a directly gazing face for 1000 ms and then a screen containing a visual analogue rating scale (VAS) asking “*How trustworthy is this person?*” At this point, a cursor was visible on the screen and participants used the mouse to click along the scale at the point that represented how trustworthy they judged that person to be. The extreme left of the scale was labelled *“Very Untrustworthy”* and the extreme right of the scale was labelled *“Very Trustworthy”*. The centre of the screen therefore represented neutral. When participants clicked on the scale, the computer recorded a rating between -100 and +100. The order of face identity on each trial was randomised.

### Passive Viewing Phases

Immediately before and after the cueing phase participants completed the passive viewing phases. In these phases, participants pressed space to initiate each trial. A fixation cross was presented for 500 ms followed by a face for 750 ms. After the face disappeared, participants were presented with a "*Please Relax*" screen for 1000 ms. There were 192 trials in total. Each face was intended to be repeated 6 times in each phase. However, randomisation was repeated after every 6 faces had been presented as opposed to 16. This only introduced slight variability into the number of times each face was presented and did not differ significantly between conditions (see Table 3).

**Gaze-Cueing Phase**

For half of the participants, faces in group A were designated congruent and would consistently look towards targets whereas faces in group B were incongruent and would consistently look away from targets. For the other half of participants, the contingencies were reversed. Participants initiated each trial with the space bar, a fixation cross appeared for 1500 ms followed by a directly gazing face for 1500 ms. The face then changed gaze direction and remained for 500 ms after which an object appeared to the left or right hand side of the face and disappeared as soon as a response was made or until 3000 ms elapsed. When a response was made, the object disappeared and the face gazed directly again for 2000 ms (see Figure 1). At the end of the trial, participants saw a screen saying "*Please Relax*" for 1000 ms. A 500 ms stimulus onset asynchrony (SOA) between gaze-cue and target object was used to ensure that it was long enough for gaze to be most strongly encoded and produce measurable ERPs but at the same time short enough to reflexively cue attention (Friesen & Kingstone, 1998). This SOA has also been used in all other studies of identity-contingent gaze-cueing and trust (Bayliss et al., 2009; Bayliss & Tipper, 2006; Rogers et al., 2014), facilitating comparison. We did not vary SOA, as this would not allow for a sufficient number of ERP trials to be calculated in the trial period when the gaze-shift occurred. There were 5 blocks in total each comprising 32 trials. Within each block, each face was presented twice, once gazing rightward and once gazing leftward. The order of trials within each block was randomized. Objects in each category were randomly sampled without repetition apart from when in the opposite orientation. Participants were told that their task was to classify the object as quickly but as accurately as possible and that the face was irrelevant to their task. Response keys were counterbalanced. Half of the participants pressed spacebar for kitchen objects and the H button for garage objects whereas the other half did vice versa. Responses were made with the index finger on the H button and thumb on the space bar. If there was no response made within 3000 ms or if the response was incorrect, an error tone sounded for 1000 ms. Participants completed 8 practice trials beforehand with unfamiliar faces that were not used in the main experiment.

The protocol of the experiment was as follows. After giving informed consent, participants were presented with all the kitchen and garage objects and asked to classify them in order to verify that they could do the gaze-cueing task properly. Feedback was given for incorrect responses. Participants then completed the initial rating phase, washed their scalp with baby shampoo after which the EEG cap was fitted, electrode sites filled with gel and electrodes attached. Participants then completed a brief eye-movement task in which they were asked to make 20 leftward, rightward, upward and downward eye-movements and eye-blinks. This provided a clean template of ocular artefacts for later removal from the experimental data using independent components analysis (ICA). Participants then undertook the initial passive viewing phase followed by the cueing phase, end passive viewing phase and end trust rating phase. Afterwards, the EEG cap was removed; participants washed their hair and were debriefed.

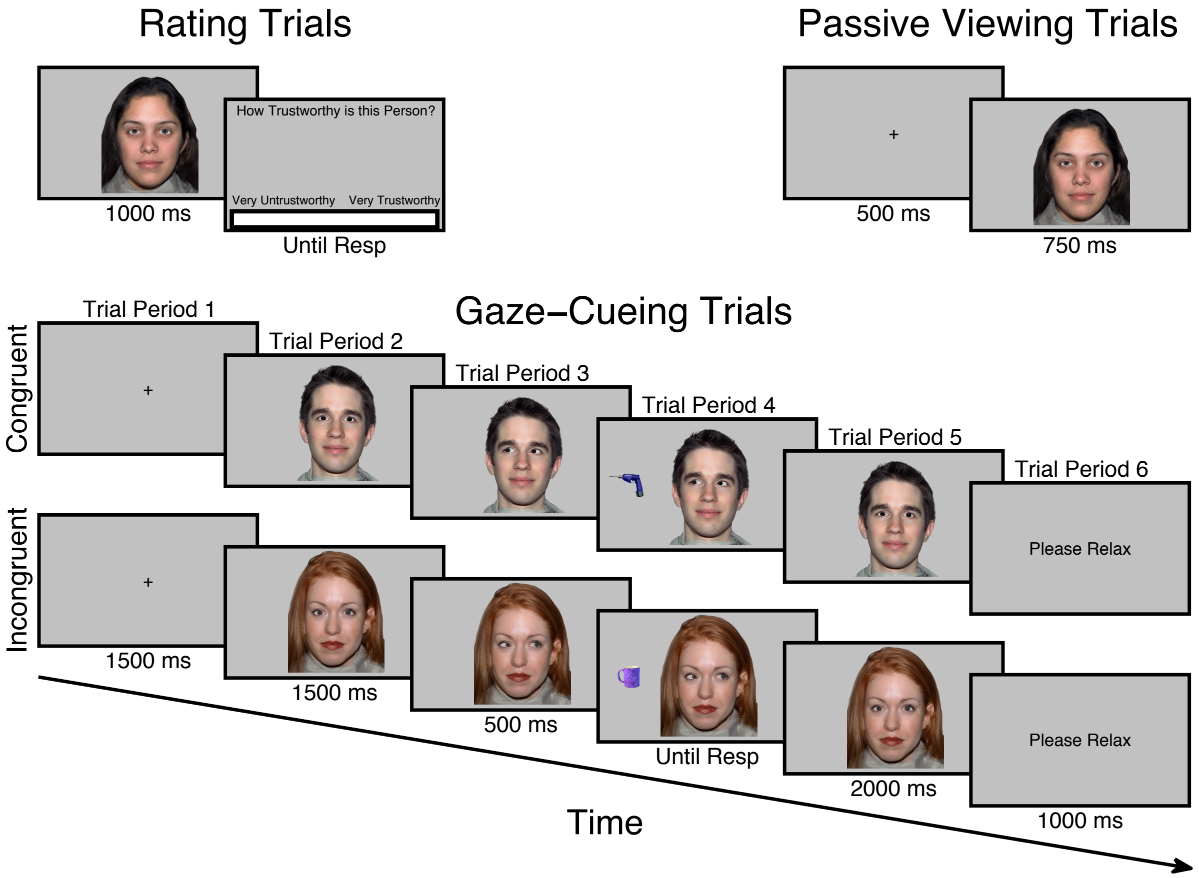


*Figure 1.* Trial procedure for trust rating, passive viewing and cueing trials. On trust rating trials before and after viewing phases, participants observed each face for 1000 ms after which a visual analogue rating scale appeared requiring them to click the point on the scale which represented how trustworthy they judged the face to be. On passive viewing trials,

which occurred immediately before and after the cueing phase, participants viewed a fixation cross for 500 ms followed by a face for 750 ms. During cueing trials, participants saw a fixation cross for 1500 ms, followed by a face looking directly for 1500 ms after which it shifted its gaze direction to the left or right for 500ms at which point a kitchen or garage target object was presented. When participants classified the object with a key-press it disappeared and the gaze returned to look towards the participant for another 2000 ms.

### EEG Data Pre-processing and Analysis

Data pre-processing was undertaken using Brain Vision Analyzer 2. The data were down sampled to 1024 Hz, filtered between 0.1-30 Hz (48 dB slope) and re-referenced to the average before being submitted to Infomax independent-components-analysis (ICA) to identify and remove eye-movements and blink artefacts. The ICA was trained on each individual's eye-movement and blink activity recorded specifically for this purpose. Errors and outliers were then removed. For the cueing phase, all trials on which incorrect target object classifications occurred or on which participants took longer than 1500 ms to respond were removed. Each trial was then visually inspected for artefacts such as excessive EMG activity (blind to conditions). Bad channels were recalculated by interpolating between neighbouring electrodes. All trials and trial periods were baseline corrected using the final 100 ms of the fixation periods. After pre-processing, all data were exported into SPM12 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) for statistical analysis, allowing for the testing of effects over all time points and scalp sites whilst correcting for the family wise error (FWE) rate with random field theory (RFT). Conditions of interest were epoched, averaged and converted into interpolated 3D images at a size of 32 × 32 voxels at each time point for cueing trials and 64 × 64 voxels at each time point for viewing trials. Images were smoothed with a 9 × 9 mm × 30 ms full-width half maximum (FWHM) Gaussian kernel. The images were entered into general linear models using a flexible factorial design. $F$-contrasts were used to test for significant effects. All effects are corrected for violations of sphericity and were only considered significant if clusters passed a cluster size threshold of $p < .0001$ RFT FWE corrected. Trust effect was entered as a covariate in all analyses. Significant effects on the scalp were followed up with distributed Bayesian 3D source reconstruction modelling using multiple sparse priors (Friston et al., 2008). This involved coregistering each individual's sensor array to a template MNI brain using standard coordinates. The source space was modelled using a "canonical" mesh of the cortical surface and the boundary element model used to account for volume conduction by the surrounding tissues whereby the cerebrospinal fluid, skull and skin are accounted for by tessellated meshes of different conductivities. Reconstruction entails computing the forward model of the lead fields from each cortical mesh vertex to the sensors and then performing inversion using the experimental data. Group inversion was employed in order to optimise the spatial covariance in reconstructed activity across subjects (Litvak & Friston, 2008; Litvak et al., 2011). For significant effects on the scalp, contrast waves of interest were computed for each subject and entered into the inversion, after which individual 3D images were generated by averaging across time windows of interest and submitted to a one-sample t-test.

**Data Screening Protocol**

For analysis of the cueing ERP data, all trials on which participants made an error or failed to respond (*M* = 2.6%, *SD* = 2.03 of trials) were removed along with artefact ridden trials (*M* = 16.95% of trials, *SD* = 10.5%) and trials with reaction times above 1500 ms (*M* = 6.8%, *SD* = 8.37% of trials) (see Table 1), a criterion typically employed in gaze-cueing studies (Langton & Bruce, 1999; Teufel, Alexis, Clayton & Davis, 2010). Paired samples t-tests showed no significant difference between congruent and incongruent conditions in the number of errors, *t* (25) = .202, *p* = .841, 95% CIs [-.44 .54], and artefacts, *t* (25) = .896, *p* = .379, 95% CIs [-.84 2.14]. However, there were significantly more outliers above 1500 ms in the incongruent condition, *t* (25) = -2.18, *p* = .039, 95% CIs [-2.2 -.06]. For trial period 4, where the duration depends upon the reaction time, extra trials were removed that were below 500 ms in duration (*M* = .55%, *SD* = 1.1%). After removal of errors, outliers and artefacts, there was no significant differences in the number of trials between incongruent and congruent conditions in trial period 4, *t* (1,26) = .649, *p* = .522, 95% CIs [-1.84 3.530], and all other trial periods, *t* (1,25) = .846, *p* = .40, 95% CIs [-1.6 3.8]. See Table 2. There was also no significant difference between incongruent and congruent conditions in terms of the number of artefact trials (*M* = 10.62%, *SD* = 9.8%) removed from the viewing analyses, *t* (25) = .597, *p* = .556, 95% CIs [-.49 .89].

For all ERP analyses, we collapsed across the factor of face gender, as this was not of primary interest to our hypotheses and also to retain a sufficient number of trials in the analysis. Using repeated-measures ANOVA on the mean number of repetitions of each face in the viewing phase, there was no effect of validity, *F*(1,25) = .151, *p* = .701, $\eta_p^2$ = .006, time, *F*(1,25) = 0.0, *p* = 1.0, $\eta_p^2$ = 0.0, or their interaction, *F*(1,25) = 1.08, *p* =.309, $\eta_p^2$ = .041. Thus, the effects would be similar to that of removing trials due to artefacts (see Table 3).

Trials with errors or reaction times exceeding 1500 ms or two standard deviations above or below each participants mean (*M* = 4.6%, *SD* = 1.5) were removed from the reaction time analyses (in accordance with Bayliss et al. (2009), Bayliss and Tipper (2006), and Rogers et al. (2014)). There were no significant differences between incongruent and congruent conditions in the number of outliers exceeding 2 standard deviations of each participants mean, *t* (25) = -1.8, *p* = .087, 95% CIs [-1.45 .106]. In order to divide participants into those showing trust effects and those who did not, each participant's mean rating change score (end – beginning) for congruent and incongruent faces were calculated and graphed. Those who showed a larger negative change to incongruent compared to congruent, without large overlap in the standard error bars, were classified as showing a trust effect (referred to in between-subjects analyses as the Trust Effect group), whereas those who showed no difference or changes in the opposite direction were classified as not showing a trust effect (referred to as the No-Trust Effect group). This was then entered into the reaction time and EEG analyses of variance (ANOVAs) as a between-subjects factor/ covariate. Overall, 17 participants exhibited a trust effect whereas 9 participants did not.

Table 1. *Means and standard deviations of the percentage of errors, outliers and artefacts removed from cueing analyses across conditions.*

| | Errors | Outliers above 1500 ms | Outliers above 2 SDs | Artefacts |
|---|---|---|---|---|
| Congruent | 1.4(1.28) | 2.07(2.9) | 1.95(0.91) | 8.9(5.6) |
| Incongruent | 1.4(1.07) | 3.17(4.4) | 2.6(1.3) | 8.2(5.6) |

Table 2. *Means and standard deviations of numbers of trials in cueing ERP analyses after removal of errors, outliers and artefacts.*

| | Block 1 | Block 2 | Block 3 | Block 4 | Block 5 |
|---|---|---|---|---|---|
| Congruent | 11.77(2.07) | 11.39(3.37) | 11.70(3.33) | 12.08(2.33) | 11.96(3.07) |
| Incongruent | 11.65(3.51) | 11.77(2.94) | 11.54(2.98) | 11.15(2.87) | 11.69(2.94) |

Table 3. *Means and standard deviations of the number of repetitions and number of trials in each condition of the viewing analyses after removal of artefacts.*

| | Beginning | | End | |
|---|---|---|---|---|
| | Congruent | Incongruent | Congruent | Incongruent |
| Repetitions Mean | 5.9(.54) | 6.15(.54) | 6.03(.54) | 5.97(.54) |
| Number of Trials Mean | 42.58(6) | 44.23(6.3) | 42.89(7.4) | 41.92(6.3) |

## Results

### Evaluations of Trustworthiness

The ratings were analysed using a 2 × 2 × 2 repeated-measures ANOVA with factors of time of rating, face gender and validity. There was a significant main effect of validity, $F$ (1,25) = 15.021, $p$ = .001, $\eta_p^2$ = .375, qualified by a significant time × validity interaction, $F$ (1,25) = 16.749, $p$ < .0001, $\eta_p^2$ = .401. This is due to more negative ratings for incongruent faces ($M$ = - 22.29, $SD$ = 7.03) compared to congruent faces ($M$ = 23.09, $SD$ = 5.88) in the final rating phase (see Figure 2). There was also a significant time × face gender interaction, $F$ (1,25) = 4.256, $p$ = .050, $\eta_p^2$ = .145. This interaction is due to an overall more negative change in ratings for female ($M$ = -9.9, $SD$ = 4.4) compared to male faces ($M$ = -.38, $SD$ = 3.98). No other effects reached significance. In order to formally identify the source of the main effects and interactions described above, separate validity × face gender ANOVAs were run on the beginning and end ratings. These analyses showed that, whereas at the beginning there was no significant effects of validity, $F$ (1,25) = .000, $p$ = .985, $\eta_p^2$ = .000, face gender, $F$ (1,25) = .049, $p$ = .826, $\eta_p^2$ = .002, or their interaction, $F$ (1,25) = 2.6, $p$ = .119, $\eta_p^2$ = .094, at the end rating, there was a significant effect of validity, $F$ (1,25) = 16.66, $p$ < .0001, $\eta_p^2$ = .40, a significant effect of face gender, $F$ (1,25) = 4.39, $p$ = .046, $\eta_p^2$ = .149, but no interaction, $F$ (1,25) = .024, $p$ = .879, $\eta_p^2$ = .001.

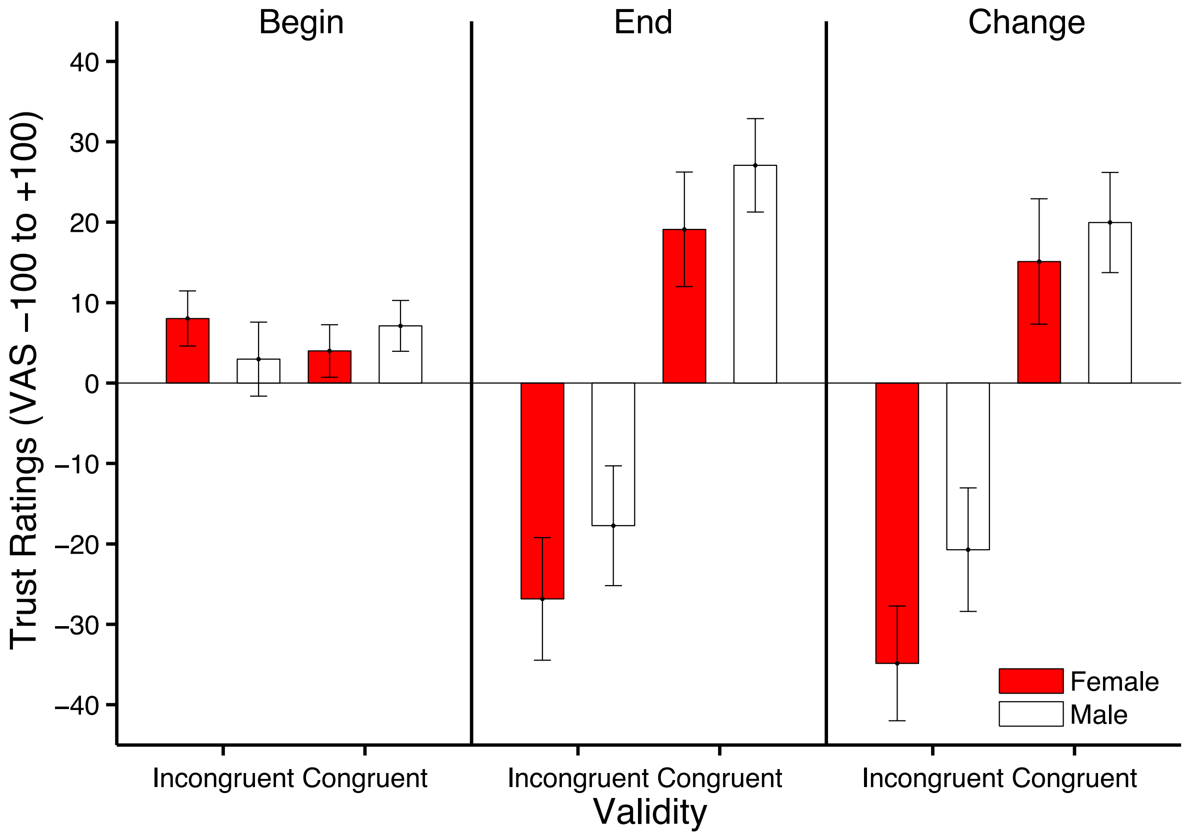


*Figure 2.* Mean trustworthiness ratings at the beginning (left) and end (middle) as well as the change in ratings from beginning to end (right) in experiment 1. Error bars show +/- 1 standard error of the mean.

### Gaze-Cueing Reaction Times

Reaction times during the cueing phase were submitted to a 2 × 2 × 2 × 5 mixed analysis of variance (ANOVA) with the within-subjects factors of face gender, validity and block and the between-subjects factor of trust effect. There was a significant main effect of validity, $F$ (1,24) = 20.02, $p < .0001$, $\eta_p^2 = .455$, demonstrating a cueing effect due to slower reaction times for incongruent ($M$ = 890.8, $SEM$ = 28.2) compared to congruent faces ($M$ = 832.9, $SEM$ = 22.468) (see Figure 3). There was also a significant main effect of block, $F$ (4, 96) = 36.5, $p < .0001$, $\eta_p^2 = .603$, and a significant linear trend for block, $F$ (1, 24) = 58.6, $p < .0001$, $\eta_p^2 = .701$, demonstrating a general decrease in reaction times as participants become more practiced. However, there was no validity × block interaction, $F$ (4, 96) = .538, $p$ = .708, $\eta_p^2 = .022$, showing that the effects of validity remained constant throughout the experiment. There was also no interaction between trust effect and validity, $F$ (1, 25) = .22, $p$ = .643, $\eta_p^2 = .009$, or trust effect, validity and block, $F$ (4, 96) = .1.31, $p$ = .273, $\eta_p^2 = .052$. However, there was a significant main effect of trust effect, $F$ (1, 24) = 8.03, $p$ = .009, $\eta_p^2$ = .251, demonstrating that participants who showed a trust effect in the ratings were generally faster at responding (*M* = 791.9, *SEM* = 29.03) than those who did not show a trust effect (*M* = 931.72, *SEM* = 39.89). No other effects reached significance.

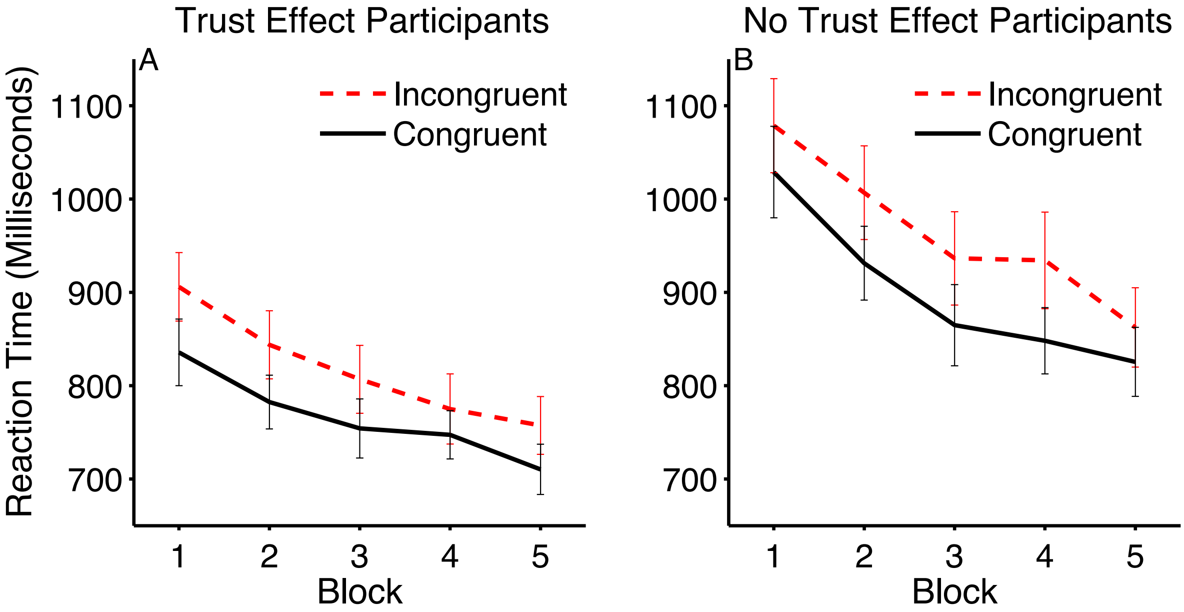


*Figure 3.* Mean reaction times by block and validity for trust effect participants (left) and no trust effect participants (right) in experiment 1. Error bars show +/- 1 standard error of the mean.

**Passive Viewing Phase ERPs**

For analysis of the viewing data, contrasts of interest were time (before/ after cueing) and validity × time. To begin with, all effects were tested at an uncorrected threshold of *F* (1,192) = 6.77, *p* < .01. There were no significant main effects or interactions apart from a strong main effect of time peaking at approximately 250 ms. This contrast was subsequently voxelwise thresholded at *F* (1,192) = 15.79, *p* < .0001, uncorrected. The effect of time was evident as two significant clusters of activity on separate posterior occipitotemporal and frontocentral electrodes between 200-300 ms, peaking at 248 ms, *F* (1,192) = 40.25, *p* < .0001, *k* = 43,386, for the former and 242 ms for the latter, *F*(1,192) = 32.58, *p* < .001, *k* = 36,790. Figure 4 shows these clusters of significant activity and illustrates the waveforms for the electrodes nearest peak voxels in the anterior (electrode C12) and posterior clusters (electrode B8). The effect is characterised by a larger negative deflection at the end compared to the beginning on posterior sites and a larger positive deflection to end compared to beginning on anterior sites. The difference between clusters reflects the typical dipolar distribution of the source of the activity and conforms to the previously reported N250 component (Joyce & Rossion, 2005; Pierce et al., 2011; Schweinberger et al., 2002; 2004;

Tanaka et al., 2006). The graphs also show the P1, N170 and P200 visual evoked components on posterior electrodes and the N1, VPP and N200 inverted counterparts on frontocentral electrodes. However, no significant effects within the time-windows of the P1, N1, N170 and VPP were observed even at low uncorrected thresholds of $p < .05$.

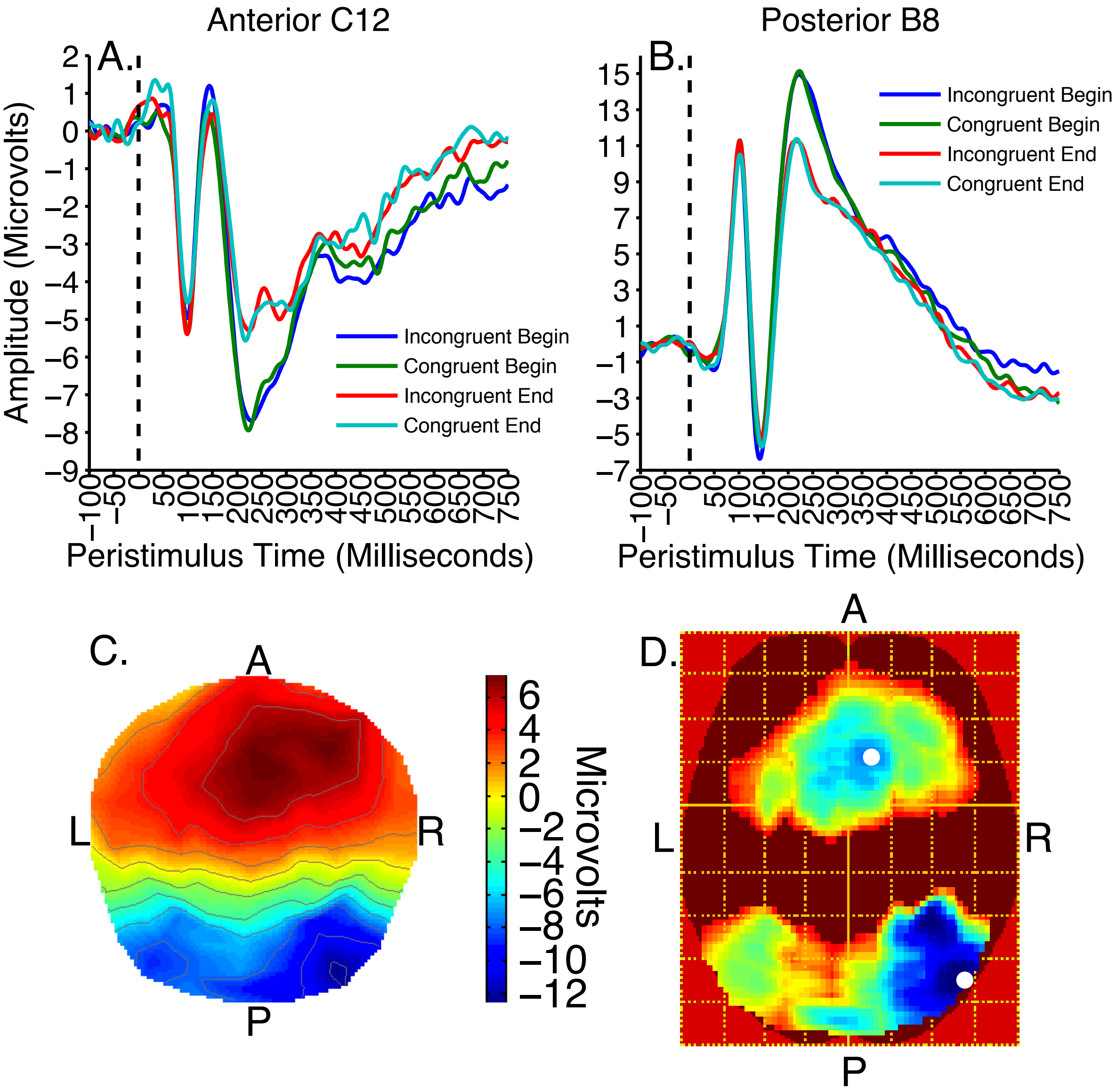


*Figure 4.* Viewing trials. Top panels illustrate time courses of activity, showing the effect of time at 250 ms on anterior cluster peak electrode C12 (A) and posterior peak electrode B8 (B). Bottom panels show the scalp distribution of activity for the effect of time at 250 ms (C), and the *F*-map of significant voxels (D), thresholded with a voxelwise height threshold of $F = 15.79$, $p < .0001$, uncorrected. Both clusters were clustersize significant at $p < .0001$ RFT FWEC. Colder colours indicate higher *F*-values. White spots denote peak locations.

## 3D Source Reconstruction

Multiple sparse priors (MSP) were used to model the source of the effect of time at 250 ms during the passive viewing phases. No smoothing was used to preserve the spatial specificity of effects to the cortical surface. This revealed two highly significant bilateral clusters of activity on the anterior middle temporal gyrus (see Figure 5). The effect appeared to be almost symmetrical across hemispheres in both location and magnitude (see Table 4). One peak in each cluster was identified, suggesting all voxels had equal strength.

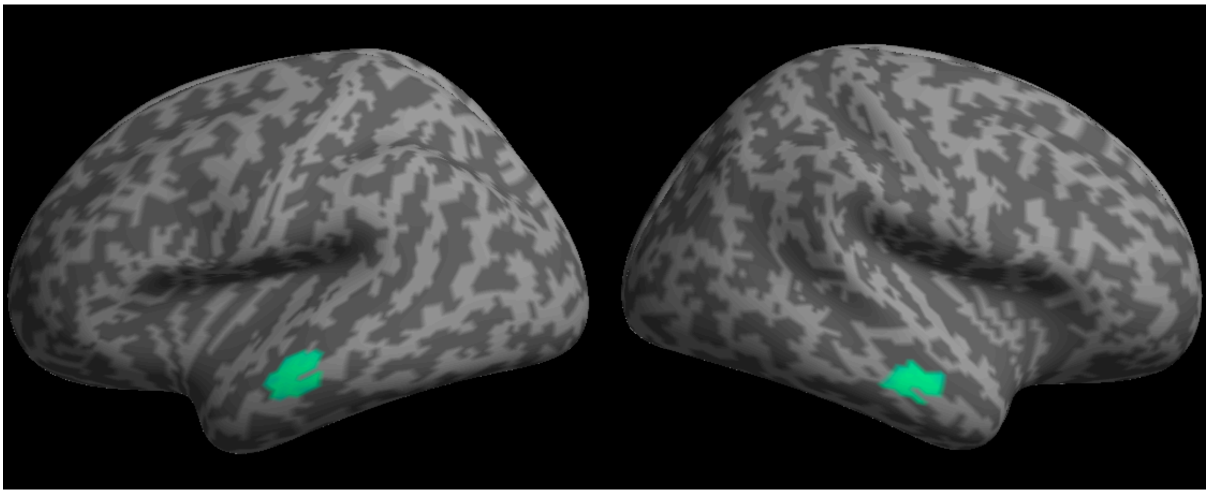


*Figure 5.* MSP localisation of the main effect of time at 250 ms rendered onto a flattened image of the cortical surface (voxelwise thresholded at $p < .001$ RFT FWEC).

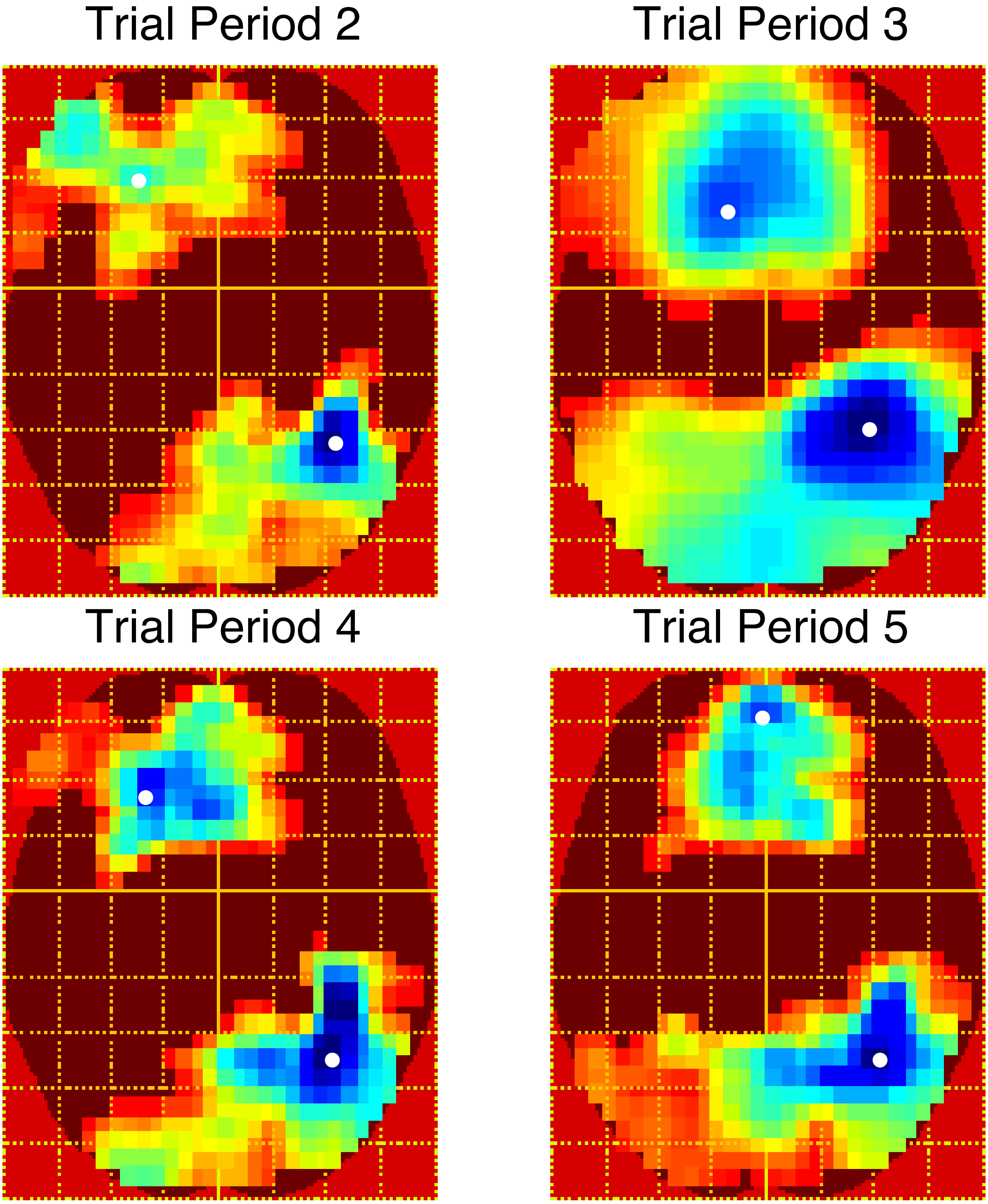
Table 4. *MNI coordinates, size, and significance of sources of the time effect at 250ms (p<.001, RFT FWEC).*

| Location | Number of voxels | Peak significance (FWEC) | MNI Coordinates of Peak | | |
|---|---|---|---|---|---|
| | | | X | Y | Z |
| R Middle Temporal Gyrus | 178 | $P < .0001$ | -64 | -18 | -10 |
| L Middle Temporal Gyrus | 173 | $P < .0001$ | 64 | -20 | -16 |

## Gaze-Cueing Phase ERPs

In the gaze-cueing phase, each trial period was analysed separately and all contrasts were voxelwise thresholded at $F(1,480) = 6.69$, $p < .01$, uncorrected. Contrasts of interest were those pertaining to validity and block × validity. We investigated block × validity with linear contrast weights such that from blocks 1 to 5 congruent faces were weighted as 2, 1, 0, -1 and -2, whereas the sign of the weights were reversed for incongruent faces. Although there were no main effects of validity, there was a significant linear block × validity interaction which emerged in the final 500 ms of trial period 2 and remained almost constant throughout the other trial periods. The distribution of activity was similar across trial periods and was evident as two large clusters of activity on separate frontal and parieto-occipital electrodes which were opposite in polarity resembling a typical dipolar pattern (see Figure 6).

Figure 7 shows the difference waves (incongruent - congruent) between validity conditions across blocks and trial periods on the electrodes nearest peak voxels in both clusters. The LPP is typically measured as an enhanced positivity over parietal sites (Schupp et al., 2000). Therefore, the patterns of responses are consistent with the interpretation of a gradual increase in the LPP to incongruent faces across blocks despite the polarity of the effect being reversed on anterior electrodes. On posterior electrodes in trial periods 2, 3 and 5, the effect appears to be due to larger LPPs to congruent faces in the first two blocks where after the LPP flips and increases for incongruent faces in blocks 3, 4 and 5. Trial period 4 shows a slightly different pattern and appears to be due to block 3 being larger than blocks 1 and 2 and block 5 being larger than block 2 and 4. The differences between trial periods 2, 3, and 5 and trial period 4 is likely due to the extra trials removed, varying trial lengths and differing processes involved in the latter. Also, this is the time point at which a lateral eye-movement is made when categorising the object and thus is more susceptible to distortion by these eye-movements and by their removal with ICA. Table 5 shows the peak times, *F*-values and extent of the effects in voxel size across trial periods and clusters. There was also no main effect of validity in trial periods 2 and 3 when block 1 was excluded from the contrast. This contrast was based on the rationale that because trial periods 2 and 3 occur before the gaze-cue, no information about congruency would be available on the first trial of the experiment. Throughout both the viewing and cueing analyses, trust effect was entered as a covariate. However, no significant interactions between trust effect and validity or trust effect, validity and block was observed even at the lowest uncorrected threshold of $p < .05$.



Trial Period 2

Trial Period 3

Trial Period 4

Trial Period 5

*Figure 6.* *F*-maps for the linear block × validity interaction across trial periods. All maps are voxelwise thresholded at $F$ (1,480) = 6.69, $p < .01$, uncorrected. However, all clusters were significant at $p < .0001$ RFT FWEC. Colder colours indicate higher *F*-values. White spots show peak voxels.

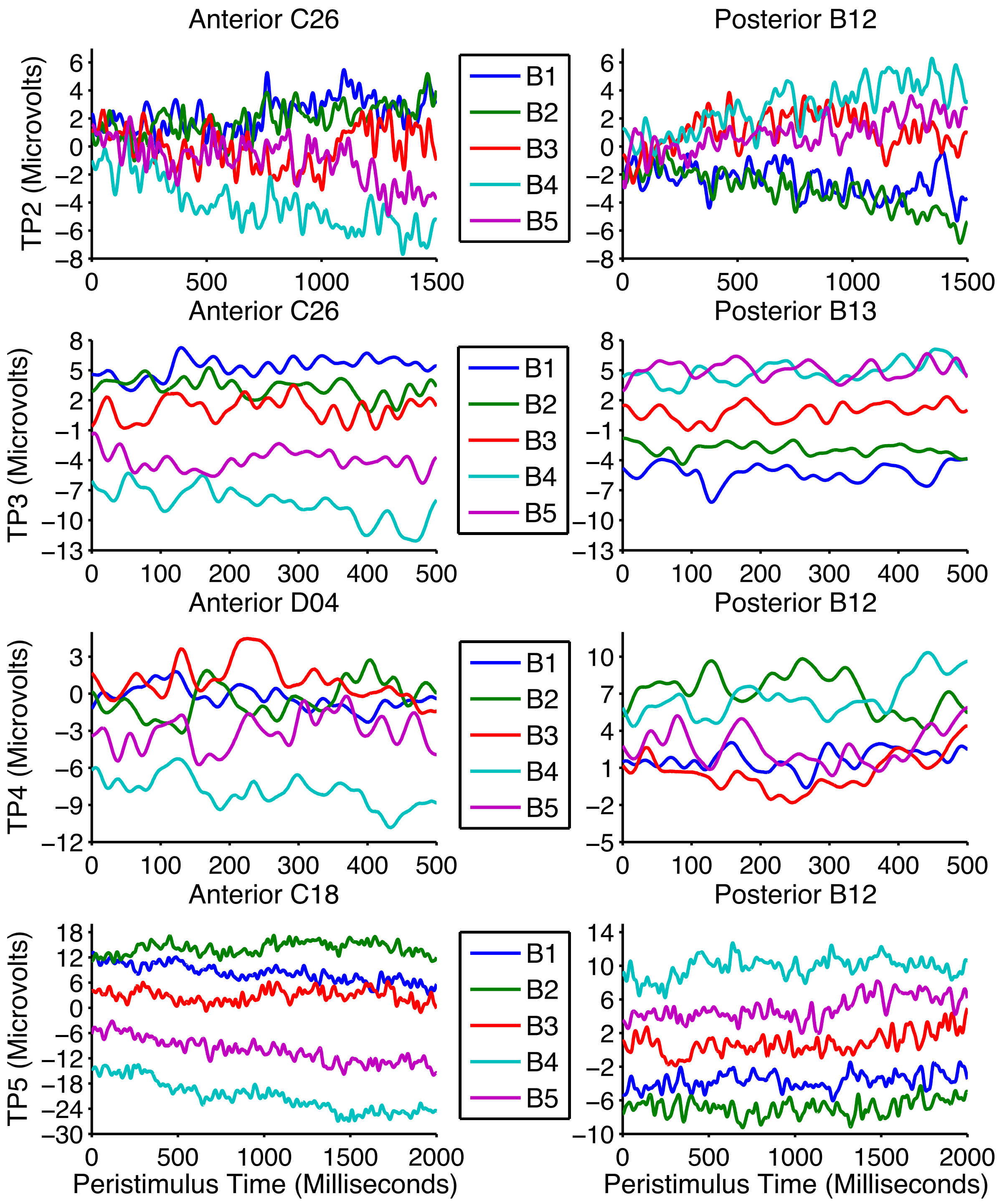


*Figure 7.* Difference waves between validity conditions (incongruent-congruent) across blocks shown separately for each trial period on the peak anterior cluster (left panels) and posterior cluster (right panels) electrodes.

Table 5. *Peak times, F-values and size of significant clusters across trial periods.*

| Trial Period | Anterior Cluster |  |  | Posterior Cluster |  |  |
|---|---|---|---|---|---|---|
|  | Peak (ms) | Peak *F*-Value | Size (Voxels) | Peak (ms) | Peak *F*-Value | Size (Voxels) |
| TP2 | 1454 | 14.11 | 41240 | 1160 | 18.82 | 40182 |
| TP3 | 132 | 24.86 | 101354 | 443 | 30.88 | 115544 |
| TP4 | 38 | 15.19 | 45662 | 225 | 17.70 | 48962 |
| TP5 | 687 | 16.8 | 171374 | 1485 | 20.37 | 279238 |

## Discussion

In this experiment we used EEG to examine electrophysiological correlates of emotion during learning of trust from identity-contingent gaze-cues. As in previous studies, reaction times to identify targets were slower when faces gazed away from the target, incongruently, compared to when gazing towards the target, congruently (Driver et al., 1999; Friesen & Kingstone, 1998). Despite the fact that the identity and gaze direction of both incongruent and congruent faces predicted target location, there was still a cueing effect in block 5. This is highly consistent with the original study of Driver et al. (1999), which found gaze-cueing effects even when participants knew the face was four times more likely to gaze in the opposite direction to the location of the target. It is also consistent with research showing gaze-cueing effects are immune to manipulations such as backward masking, working memory load, perceptual load, binocular rivalry and rapid serial visual presentation (Law et al., 2010; Sato et al., 2007; Xu et al., 2011).

In addition to standard gaze-cueing effects we replicated the effects of these cues on trust judgements as incongruent faces became judged less trustworthy than congruent faces (Bayliss et al., 2009; Bayliss & Tipper, 2006; Rogers et al., 2014). However, the current findings go beyond previous work in two ways. First, we have employed an initial baseline measure of trust allowing us to observe the direction of the effect. It was possible that congruent faces could increase in trustworthiness with no effect for incongruent faces, incongruent faces could decrease in trustworthiness with no effect for congruent faces, or that both congruent and incongruent faces would show changes in trustworthiness in opposite directions. The data support the latter of these alternatives. The second new finding concerns stimulus gender differences. The positive change in trust ratings in response to congruent cues were greatest for male faces whereas the negative change in response to incongruent cues were larger for female faces resulting in an overall more negative change in ratings for females compared to males as evidenced by the time × gender interaction. During gaze-cueing, the learning of trust was paralleled by the late positive potential component (LPP), related to emotion processing, which increased to incongruent faces across blocks after an initial response to congruent faces in the early blocks. This effect is highly consistent with previous research showing that the LPP is modulated by both positive and negatively valenced stimuli (Cuthbert et al., 2000; Schupp et al., 2000) and is related to emotional learning (Franken, Huijding, Nijs & Van Strien, 2011; Sánchez-Nàcher, Campos-Bueno, Sitges & Montoya, 2011) and memory (Smith, Dolan & Rugg, 2004).

During gaze-cueing it is important to note that trial periods 2 and 3 are qualitatively different to periods 4 and 5. In the former trial periods 2 and 3 any differences between congruent and incongruent faces must be because of prior learning of the association between face identity and gaze congruency. This is because, in trial period 2, when the face is initially looking straight ahead, and in trial period 3, when gaze has shifted, the absence of a target means participants cannot tell whether this will be a congruent or incongruent trial. Only via retrieval of prior episodes of gaze-cueing behaviour evoked by the particular viewed face, can the validity of the face be known in advance of the target. It would therefore appear that retrieval of a face's prior gaze-cueing behaviour takes approximately 1000 ms, as it is only in the last 500 ms of trial period 2 that significant LPP effects are detected. In contrast, trial period 4 reflects the period where gaze is directed towards or away from the target, so there is an explicit signal as to whether the face deceives or not; while trial period 5 reflects the situation where review of the previous congruent or incongruent face can take place.

The significant change in EEG activity between blocks 2 and 3 as the experiment progresses would appear to be the time when a qualitative change in the representation of the faces takes place. That is, the time when the salience of the incongruently gazing face that is misleading and deceiving the participant is represented. However, the effect appeared not to be perfectly linear. In all trial periods the largest LPP difference to incongruent was observed in block 4 and in trial period 5, after the gaze-cue occurred, the largest response to congruent faces occurred in block 2. This suggests a role for learning and habituation. Thus, learning about congruent faces may peak in block 2 before habituating and increasing to incongruent faces until block 4 after which the response to incongruent faces also begins to habituate. Such habituation during the final block may explain why no effects of validity were observed in the final passive viewing phase, despite a large N250 face familiarity effect.

The peak of the cueing LPP effects appeared to localise to the right posterior hemisphere. This is highly consistent with both the emotion and face processing literature. The core face processing areas of the occipital and temporal lobes (the occipital face area, fusiform face area and superior temporal sulcus) are right hemisphere dominant (Kanwisher & Barton, 2011) and the right hemisphere has been proposed to be specialised for emotion processing (Silberman & Weingartner, 1986) or processing negative emotions (Harmon-Jones et al., 2010). For both reasons, the LPP effects related to faces may be more right hemisphere distributed, when displayed with an average reference. The LPP has been proposed to be due to enhanced attention to motivating stimuli (Hajcak et al., 2010) and this may be reflected in increased processing in category specific brain regions, shifting the scalp distribution of peak activity to the right hemisphere. The parietal maximum is also consistent with neurophysiological findings that responses of neurons in the lateral intraparietal area (LIP), which encode the importance of where to look in retinotopic space (Bisley & Goldberg, 2010), are larger for social stimuli which the monkey values (Klein et al., 2008), including gaze-cues which signal stimuli of value (Shepherd et al., 2009). This value signal may originate from the orbitofrontal cortex (Watson & Platt, 2012).

The increase to incongruent faces across blocks is in line with the notion that participants initially anticipate a pro-social, trustworthy, interaction, as shown in the initial explicit trust ratings, but that learning has to gradually occur as expectancies are repeatedly violated. Indeed, a one-sample t-test on the beginning trust ratings showed that initial trust was significantly greater than zero, $t\ (25) = 2.4$, $p = .027$, 95% CI's [.69 10.36]. Previous studies have shown that the LPP to oddball negative stimuli is larger than to oddball positive stimuli among more frequently presented neutral stimuli during evaluation, particularly in younger adults (Ito et al., 1998; Hilgard et al., 2014; Wood & Kisley, 2006). In addition, untrustworthy faces elicit a larger LPP than trustworthy faces (Marzi et al., 2012; Yang et al., 2011). Our findings also suggest that negative, untrustworthy, incongruent cues are given greater weight than the positive, trustworthy, congruent cues. We repeated our analysis using the linear contrast weights separate for the block effect for incongruent and congruent faces. All clusters in all trial periods were significant for incongruent but not congruent faces. This negativity bias for incongruent gaze is also consistent with a study, which showed that the pupil size of participants, which is related to emotional arousal (Bradley et al., 2008), was larger during non-joint attention compared to joint-attention trials (Schilbach et al., 2010).

The LPP is reduced when attention is cued to a non-arousing portion of unpleasant pictures (Dunning & Hajcak, 2009; Hajcak, Dunning & Foti, 2009). Therefore, the linear effect may be due to increased attentional salience by emotion to the negatively judged incongruent faces as blocks progress. In turn, this may facilitate the learning of face valence to produce changes in trust judgements and may explain the increased feelings of familiarity for incongruent compared to congruent faces (Bayliss & Tipper, 2006). The LPP to unpleasant images is also reduced after cognitive reappraisal, where the image is reinterpreted in a less negative way (Foti & Hajcak, 2008; Hajcak & Nieuwenhuis, 2006; Moran et al., 2013). The interaction of the LPP with cognition may relate the LPP to appraisal processes at rating. However, we did not find any interactions with trust effect, possibly due to small numbers of participants not showing trust effects.

In the passive viewing trials, faces were presented at the start of the experiment to provide a baseline measure of face-related ERPs, and then at the end of the experiment the faces were again passively viewed in an attempt to detect whether the faces that had consistently looked at targets, congruently, could be discriminated from those consistently looking away, incongruently. The results from the passive viewing conditions confirmed previous findings concerning face repetition/ familiarity. We found a strong N250 familiarity effect, where the ERP signal around 250 ms on posterior occipitotemporal sensors was significantly changed from first viewing of faces relative to viewing at the end of the experiment after numerous exposures. Interestingly, exploratory source localisation using multiple sparse priors clearly identified bilateral anterior temporal (ATL) cortical sources for the N250. This is in contrast to earlier studies suggesting a more posterior fusiform gyrus source (Schweinberger, Kaufmann, Moratti, Keil & Burton, 2007; Schweinberger et al., 2002) using the brain electrical source analysis (BESA) approach. However, our findings are highly consistent with fMRI and single-unit recording studies in both macaque monkeys and humans (Freiwald & Tsao, 2010; Tsao, Moeller & Freiwald, 2008). For example, in humans, multivoxel pattern classification of fMRI data has found voxels in ATL that can reliably differentiate between different faces (Kriegeskorte et al., 2007) and activity in the same region of ATL correlates with behavioural measures of face recognition performance (Nasr & Tootell, 2012). Although our source localisation identified bilateral activations, the effect on the scalp was strongest on the right which is consistent with the studies described above, which showed that ATL face patches are bilateral but stronger on the right.

However, our main concern was to identify the neural signal for face-trust learning. We know from participants' explicit reports that the gaze-cueing procedure significantly changed their trust ratings of the faces. Yet in the analysis of the ERP response during passive viewing of the faces we found no evidence for such discrimination. We believe that the contrast between behaviour and neural activity is because the faces in the passive viewing procedure were presented for a relatively brief period of 750 ms. The analysis of the gaze-cueing procedure suggests that the statistically significant discrimination of different trust assessments emerges after 1000 ms. Furthermore, as noted above, there may have been habituation processes which could have reduced detection of gaze-cue validity towards the end of the experiment when participants passively viewed the faces with no explicit task. Alternatively, the N250 and LPP in the end-viewing phase may have been equally sensitive to the affective qualities of both congruent and incongruent faces, but the reason no difference was observed during passive viewing may be due to a bivalent response profile. Another potential reason why effects in the LPP were limited to within the gaze-cueing task may be because of the proposed relationship between attention and the LPP. The brain activity related to attentional processing required to perform the gaze-cueing task may make the threshold lower for emotion to also influence attention and the LPP. Dissociation between the neural substrates for face recognition and emotional responses to familiar faces is consistent with lesion studies. Tranel, Damasio and Damasio (1995) showed that whereas patients with damage to the ventral occipitotemporal face processing pathways show impairment in face recognition they still show evidence of emotional responses to familiar others, as measured with skin conductance (SCR). In contrast, patients with ventromedial frontal lesions recognise familiar faces but do not show SCR responses.

Whilst experiment 1 has provided evidence consistent with the notion that emotion mediates the learning of trust from identity-contingent gaze-cues, it must be acknowledged that the veracity of our conclusions are limited by the fact that the congruent and incongruent faces were confounded with large and small pupil size, respectively. It is possible that the effects in the ratings and EEG are not only influenced by emotional responses to pupil size, but these perceptual cues may be associated with the faces in memory, acting as distinctive retrieval cues. Therefore, in the subsequent three experiments, we omit the pupil size manipulation to ascertain if effects of identity-contingent gaze-cues on emotion and trust are still obtained, in which case we can more confidently attribute the effects in experiment 1 to the gaze-cueing manipulation. As we will see, this appears to be the case. It is also noteworthy, that it is large pupil size that typically elicits larger neural activity associated with emotion (e.g. in the amygdala) than small pupil size (Amemiya & Ohtomo, 2012; Demos, Kelley, Ryan, Davis & Whalen, 2008). Yet, the effects in experiment 1 were driven by incongruent faces, all of which had small pupil size.

In conclusion, here we presented unique data from EEG for evidence for the role of emotion in the learning of trustworthiness from gaze-cues. We found that the emotion-related late positive potential (LPP) increased across blocks for incongruent compared to congruent faces possibly reflecting increased emotion, attention to and learning about, faces that deceive. The neural signature for this encoding of deceptive incongruent gaze-cueing behaviour appeared to emerge between blocks 2 and 3. The discrimination of congruent and incongruent faces in early periods of the trial reflects retrieval of prior gaze-cueing behaviour and this retrieval process takes approximately 1000 ms to emerge.

# Chapter 3

## Experiment 2: Facial electromyography (EMG) evidence for the role of embodied emotional states in the learning of trustworthiness from eye-gaze cues

Experiment 1 identified a role for emotion in the learning of trust from identity-contingent gaze-cues. The emotion related LPP increased across blocks to incongruent compared to congruent faces suggesting a negativity bias for untrustworthy faces. In experiment 2, we recorded facial EMG from the corrugator and zygomaticus muscles to examine the importance of embodied emotional states in the learning of trust. Facial EMG is a sensitive measure of the valence of emotional reactions and can provide complementary support for the findings of experiment 1. However, more importantly, facial EMG reflects somatic markers that have been proposed to be integral to emotion, emotional learning and the way in which emotion influences cognition and decision-making (Damasio, 1996).

Different patterns of EMG activity correlate with various indexes of emotion. The contraction of the corrugator is a crucial action unit for the expression of anger and distress (Ekman & Friesen, 1978). Strong evidence for the relation of corrugator EMG activity to negative emotion is demonstrated by the reliable correlations with ratings of affect in response to affective pictures. Corrugator activity increases linearly as ratings become more negative and decreases as ratings become more positive (Greenwald et al., 1989; Lang et al., 1993; Larsen et al., 2003), making it the most discriminative measure of the valence of emotional reactions. It is also correlated with amygdala responses to affective pictures (Heller et al., 2011; 2014), a brain region that is associated with emotion, facial expression and the processing of social-affective stimuli (Gothard, 2014; Mattavelli et al., 2014; Rolls, 2008).

Therefore, we predict that corrugator activity will be larger on incongruent trials, reflecting the negative emotion evoked when another person directs attention away from a relevant target object, resulting in deception or failures of joint attention. Predictions for the zygomaticus are less clear because this muscle shows a bivalent response profile, correlating with affective ratings of both positive and negative stimuli (Greenwald et al., 1989; Lang et al., 1993; Larsen et al., 2003) and is involved in both smiling and grimacing (Bradley & Lang, 2000; Ekman & Friesen, 1978). Given the strong relationship between affective ratings and corrugator activity and the hypothesised importance of embodied emotion in learning and cognition, it might be predicted that this measure may relate more clearly to participants trust ratings. Therefore, we predict that corrugator activity will more clearly differentiate participants who show trust effects from those who do not. Indeed, Neta, Norris and Whalen (2009) found that corrugator activity to surprised faces was larger in those participants who rated these faces more negatively. Experiment 2 is similar in most ways to experiment 1 except that EMG was recorded instead of EEG, there were no passive viewing phases and the pupil size of the faces during cueing were not manipulated to appear larger for congruent faces.

## Method

### Participants

Participants were comprised of 50 volunteers from Bangor University with a mean age of 22 years (*SD* = 4) and a gender split of 25 males and 25 females. Participants were mostly right-handed (*N* = 43), neurologically normal with normal or corrected-to-normal vision and received course credit for taking part.

### Stimuli and Apparatus

The face stimuli were the same 16 full colour images of 8 male and 8 female faces with facial expressions that were mildly happy used in experiment 1 except that the pupils were not manipulated to be larger for congruent faces and smaller for incongruent faces. Faces were presented centrally at a pixel resolution of 300 × 385. The target stimuli were also the same as in experiment 1 and were presented at a resolution of 175 × 175 pixels. The experiment was run on E-Prime 1.0 and displayed on a 19” Iiyama Vision-master CRT display with a screen resolution of 800 × 600 pixels in the cueing phase and 640 × 480 pixels in the rating phase.

### Design and Procedure

Before the experiment began, participants were briefed and then presented with a slide show of all the kitchen and garage target objects that would be presented in the experiment to check that they could perform the task correctly as in experiment 1. The electrodes were then mounted on the face. Areas of the skin on the forehead, left brow and cheek regions were cleaned using facial cleanser, alcohol swabs and gently exfoliated with an abrasive pad before electrode gel was applied to the electrode sites. Two 4 mm silver-chloride electrode pairs were filled with conductive gel and attached to the approximate locations of the corrugator supercilii and zygomaticus major muscles using double sided adhesive electrode cuffs. A ground electrode was attached to the forehead. In order to avoid demand characteristics associated with awareness that the EMG is recording emotional facial expressions, participants were told the cover story that the electrodes on the forehead were measuring frontal lobe EEG activity and that the electrodes on the cheek were reference electrodes in a face and object recognition ERP experiment (Fridlund & Cacioppo, 1986; Tassinary et al., 2007). Whilst the electrodes settled onto the skin, participants completed the State-Trait Anxiety Inventory (Spielberger, 1983). Participants then began the three phases of the experiment: (1) initial trust rating phase, (2) gaze-cueing task phase, (3) final trust rating phase (see Figure 8).

### Trustworthiness Rating Phases

Both beginning and end rating phases were the same as in experiment 1 (see Figure 8).

### Gaze-Cueing Phase

The gaze-cueing phase was exactly the same as in experiment 1 except that participants were given a maximum of 5000 ms to respond and the "*Please Relax*" screen was presented for 4000 ms (see Figure 8). The pupil size of the faces was also not manipulated.

### Electromyographic Recording

Electromyography was recorded using a Biopac MP100 system with two EMG100C amplifiers. All data were sampled at 2 kHz, amplified by a factor of 5000 and filtered online with a high pass filter of 10 Hz, a low pass filter of 500 Hz and a notch filter at 50 Hz. Subsequently, the data were band pass filtered between 20-400 Hz which has been shown to be the optimal bandwidth for the removal of artefacts such as eye-movements, eye-blinks and brain activity from facial EMG (Van Boxtel, 2001). As EMG data are bipolar voltages around zero, the data were full wave rectified to measure amplitude. Activity in all trial periods was divided by the mean activity between 900-1400 ms of the fixation period of each trial. This baseline correction controls for the fluctuations and noise in the EMG signal that varies over the course of the experiment. EMG activity is therefore expressed as a ratio of muscle activity with respect to baseline. The time courses of each individual's EMG activity across all trials were also inspected visually whilst blind to conditions. Separate sets of artefact trials were removed for the corrugator and zygomaticus muscles. Trials containing large inflections caused by non-expressive movements, such as yawns or coughs were removed. The data were averaged into 100 ms time bins and analysed in SPSS 20.

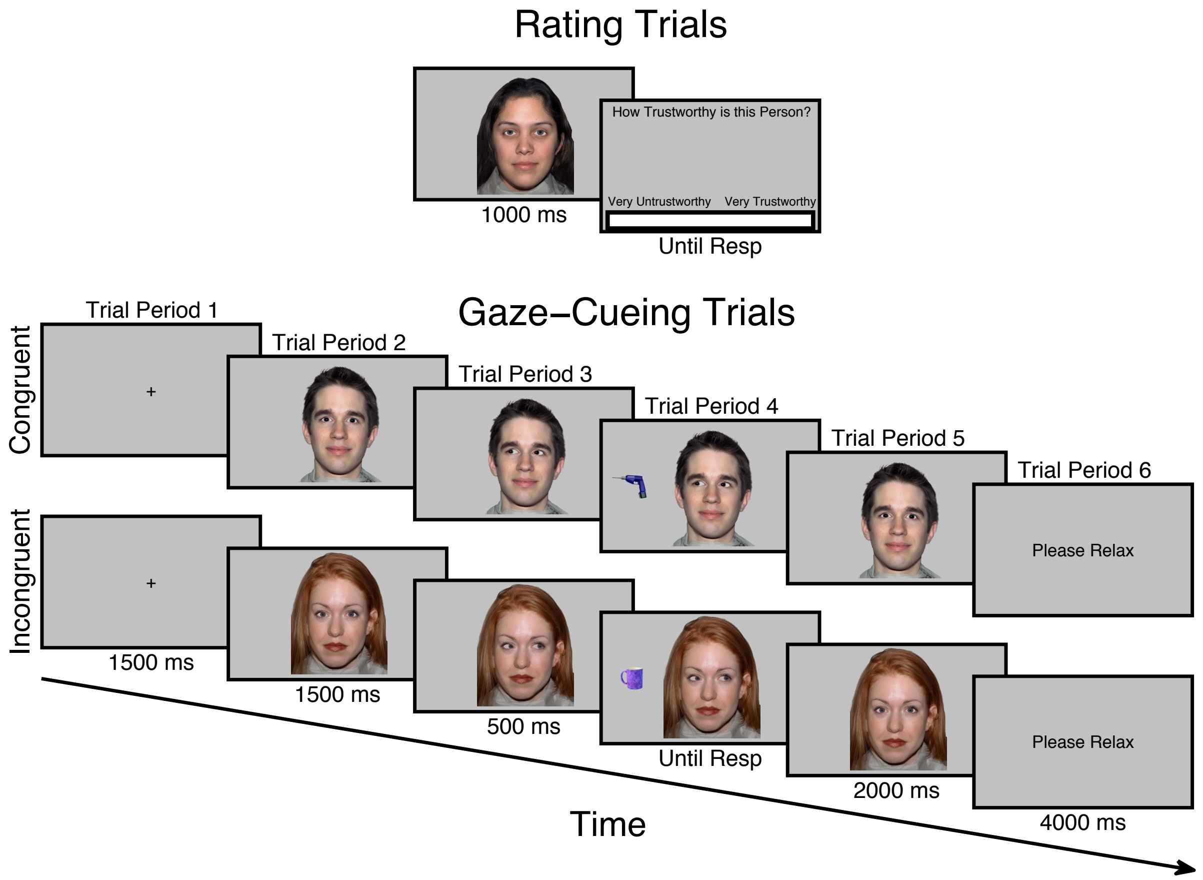


*Figure 8.* Schematic illustration of trial procedure for rating and cueing trials in experiment 2. On rating trials before and after cueing, participants observed each face for 1000 ms after which a visual analogue rating scale appeared requiring participants to click the point on the scale that represented how trustworthy they judged the face to be. During cueing trials, participants saw a fixation cross for 1500 ms, followed by a face looking directly for 1500 ms after which it changed its gaze direction and remained for 500 ms when an object appeared to the left or right hand side of the face and disappeared when the participant responded which also triggered the face to look back directly at the participant for another 2000 ms.

## Data Screening Protocol and Analysis

Table 6 shows the mean percentages of trials removed from analyses. All trials on which participants made an error or did not respond (*M* = 2.51%, *SD* = 1.95% of trials) were removed from reaction time and EMG analyses along with trials with reaction times above or below two standard deviations from each participant’s mean (*M* = 4.31%, *SD* = 1.34% of trials). In addition, any remaining reaction times larger than 1500 ms were removed (*M* = 1.95%, *SD* = 3.33% of trials). The first trial of the experiment and trials immediately succeeding error trials were removed from the EMG data to eliminate facial reactions to novelty and errors (*M* = 2.93%, *SD* = 1.65% of trials). Trials with artefacts observed were removed from the corrugator (*M* = 5.28% of trials, *SD* = 2.98%) and zygomaticus (*M* = 8.6% of trials, *SD* = 3.7%) analyses separately. Paired samples t-tests were used to test for differences in the number of errors, outliers and artefacts between incongruent and congruent conditions. There were no significant differences in outliers, *t* (49) =-1.02, *p* = .313, 95% CIs [-.71, .23], corrugator artefacts, *t* (49) =-.594 , *p* = .555, 95% CIs [-.55, .37], zygomaticus artefacts, *t* (49) = -.375 , *p* = .709, 95% CIs [-.79, .54], and errors, outliers and corrugator artefacts combined, *t* (49) = -1.85*,* *p* = .071, 95% CIs [-1.46, .063] or errors, outliers and zygomaticus artefacts combined, *t* (49) = -1.64*,* *p* = .108, 95% CIs [-1.56, .16]. Errors were significantly more frequent on incongruent trials, *t* (49) = -2.04, *p* = .047, 95% CIs [-.671, -.004] (see table 6), supporting the longer reaction times on such trials. As in experiment 1, all participants were divided into those who showed trust effects and those who did not. This was entered into the reaction time and EMG ANOVAs as a between-subjects factor. Overall, 27 participants showed trust effects whereas 23 participants did not. All analyses of between-subjects effects were reproduced with participant gender as the between-subjects factor. As no effects of participant gender were significant we refrain from reporting them here. For analysis of the EMG data in trial period 4, where the duration depends upon the reaction time, the trial period was cut down to a maximum duration of 1000 ms as this is roughly equal to the mean reaction time + 1 standard deviation.

Table 6. *Means and standard deviations of the percentage of errors, outliers and artefacts across conditions.*

| | Errors | Outliers | Corrugator Artefacts | Zygomaticus Artefacts |
|---|---|---|---|---|
| Congruent | 1.09(1.05) | 3.01(2.15) | 2.58(1.52) | 4.24(2.17) |
| Incongruent | 1.43(1.18) | 3.25(2.14) | 2.70(1.73) | 4.36(2.17) |

## Results

### Evaluations of Trustworthiness

Trust ratings were analysed using a 2 × 2 × 2 within-subjects ANOVA with factors of validity, face gender and time of rating. There was a significant effect of validity, *F* (1,49) = 11.12, *p* = .002, $\eta_p^2$ = .19, but more importantly, a significant validity × time of rating interaction, *F* (1,49) = 12.895, *p* = .001, $\eta_p^2$ = .208. As shown in Figure 9, these effects are clearly driven by the higher trust ratings for congruent (*M* = 14.108, *SEM* = 3.74) compared to incongruent (*M* = -12.45, *SEM* = 4.65) in the final rating phase. There was also a significant interaction between time and face gender owing to higher trust ratings for females (*M* = 7.585, *SEM* = 2.731) compared to males (*M* = -1.297, *SEM* = 2.813) in the initial rating phase compared to the final rating phase, *F* (1,49) = 7.352, *p* = .009, $\eta_p^2$ = .130. No other effects reached significance.

To formally identify whether the sources of the main effects and interactions described above occurred before or after cueing, 2 × 2 within-subjects ANOVAs with factors of validity and face gender were run on the beginning and end ratings separately. These analyses showed that, in the initial ratings, there was a significant effect of face gender, *F* (1,49) = 8.46, *p* = .005, $\eta_p^2$ = .147, but no significant effect of validity, *F* (1,49) = .029, *p* = .865, $\eta_p^2$ = .001, and no validity × face gender interaction, *F* (1,49) = .493, *p* = .486, $\eta_p^2$ = .01. In contrast, after cueing, there was a significant effect of validity, *F* (1,49) = 13.397, *p* = .001, $\eta_p^2$ = .215, but no significant effect of face gender, *F* (1,49) = .012, *p* = .914, $\eta_p^2$ = .000, and no validity × face gender interaction, *F* (1,49) = .071, *p* = .791, $\eta_p^2$ = .001.

The larger negative change in trust ratings for female compared to male incongruent faces and the larger positive change in trust ratings for male compared to female congruent faces, as evidenced by the time × gender interaction, is consistent with experiment 1 and the idea that learning is greater when the gaze-cues are mismatched with the kinds of social interaction expected from the visual appearance of the face. Indeed, Bayliss et al. (2009) only found trust effects with smiling and not frowning faces, where the former would be initially rated as trustworthier (Oosterhof & Todorov, 2008). This type of learning is akin to the Rescorla and Wagner (1972) learning rule, where learning is greatest when there is a discrepancy between expected and actual events.

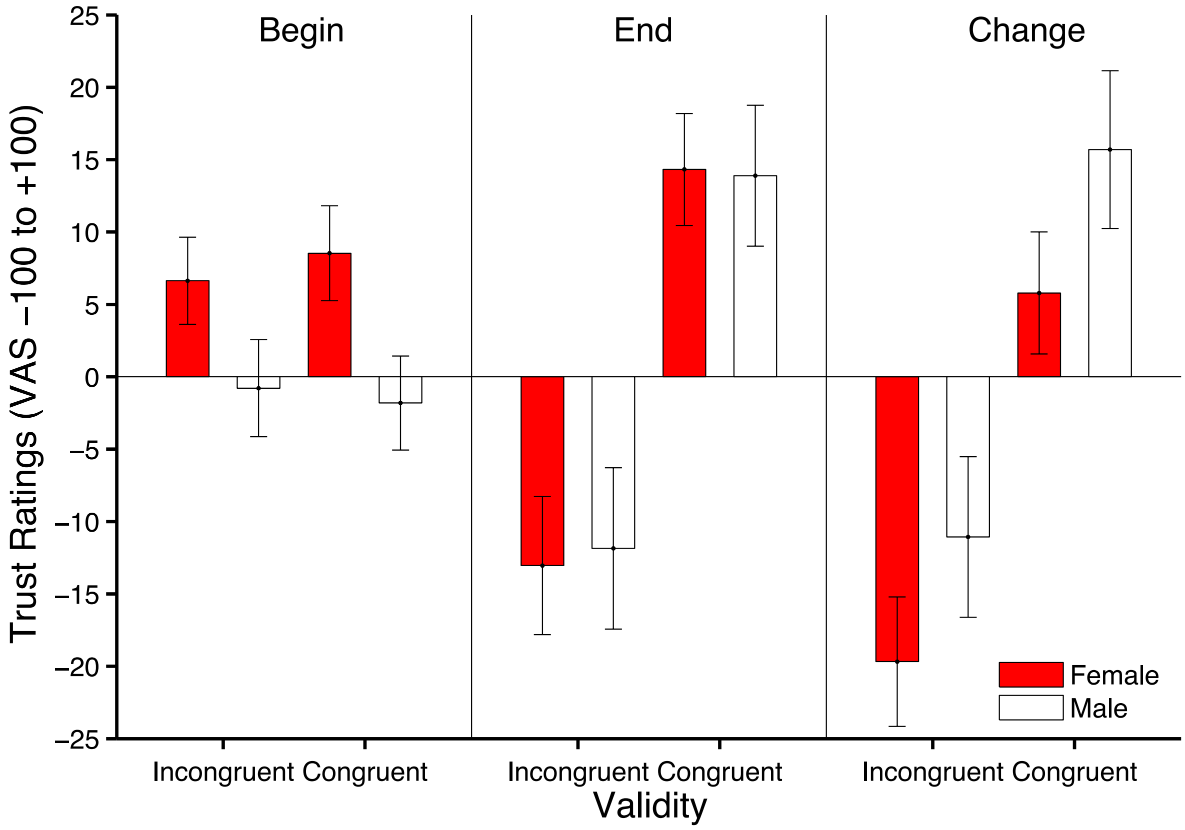


*Figure 9.* Mean trustworthiness ratings from experiment 2 by validity and face gender, given before (left) and after (middle) cueing and the change in ratings (right), computed by subtracting beginning from end ratings. Error bars show +/- 1 standard error of the mean.

**Gaze-Cueing Reaction Times**

The reaction time data were analysed using a 2 × 2 × 2 × 5 mixed design ANOVA with the between-subjects factor of trust effect and within-subjects factors of validity, face gender and block. On average, participants' reaction times were quicker on congruent trials ($M = 802.28$, $SEM = 17.68$) compared to incongruent trials ($M = 831.85$, $SEM = 16.74$) and this effect of validity was significant, $F\ (1, 48) = 35.038, p < .0001, \eta_p^2 = .422$ (see Figure 10). There was also a significant effect of block, which was due to a linear trend for a decrease in reaction times over the course of the experiment, $F\ (1,48) = 148.497, p < .0001, \eta_p^2 = .756$. However, there was no validity × block interaction, $F\ (4,192) = .840, p = .502, \eta_p^2 = .017$, suggesting that participants were not using the identity and initial gaze direction of the face to anticipate the location of the target. Importantly, there was no main effect of trust vs. no-trust group, $F\ (1,48) = 2.04, p = .160, \eta_p^2 = .041$, and no interaction between trust group and validity, $F\ (1,48) = .009, p = .926, \eta_p^2 = .000$. Separate analysis of each group confirmed significant effects of validity within both the trust group, $F\ (1,26) = 16.23, p <.0001, \eta_p^2 = .384$, and the no-trust group, $F\ (1,22) = 20.916, p <.0001, \eta_p^2 = .487$. In order to rule out the possibility that the effects of validity described above may be due to a speed-accuracy trade-off, the same 2 × 2 × 2 × 5 ANOVA was run on the reaction times weighted by proportion of correct responses in each condition. This analysis showed that the effect of validity, $F(1, 48) = 16.067$, $p < .0001$, $\eta_p^2 = .251$, and block, $F(1,48) = 95.582$, $p < .0001$, $\eta_p^2 = .668$, was still highly significant when accuracy was controlled for. No other effects reached significance.

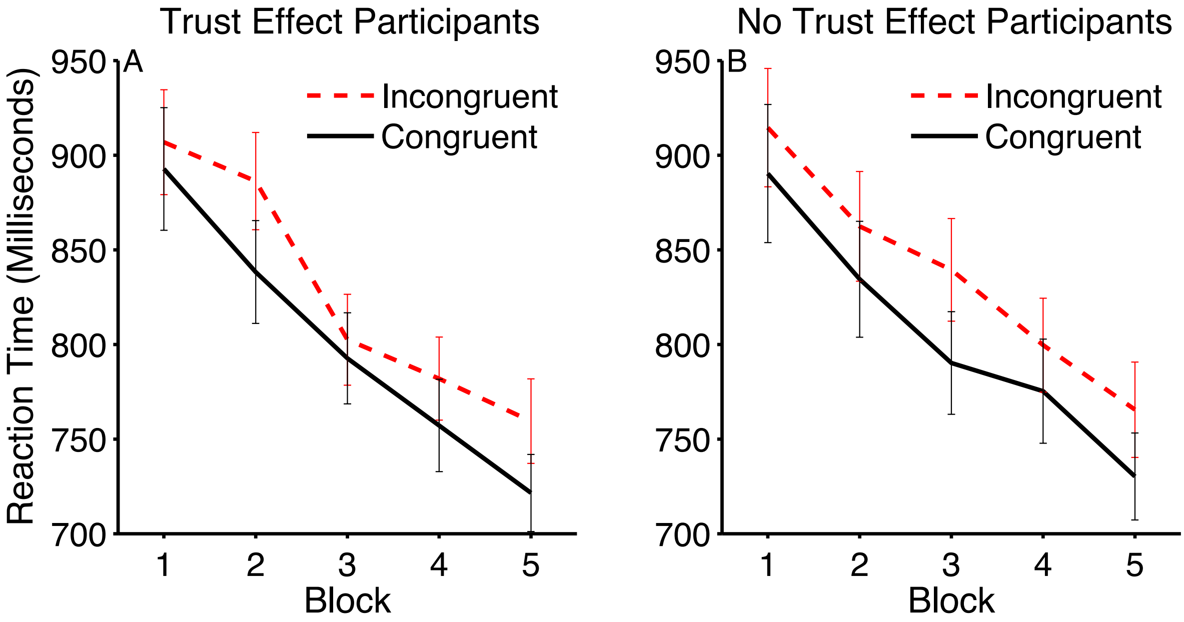


*Figure 10.* Mean reaction times on congruent and incongruent trials across blocks within participants who did (left panel) and did not show trust effects (right panel) in experiment 2. Error bars show +/- 1 standard error of the mean.

### Facial Electromyography Results

Figure 11 (corrugator) and Figure 12 (zygomaticus) show the EMG activity throughout the critical trial periods of 2 to 5. It should be noted that trial periods 2 and 3 occur before the gaze-cue and target are presented and so any contrasts between congruent and incongruent faces must be due to learning via previous exposure to the identity-contingent gaze-cues. This means that in trial periods 2 and 3 of block 1, there will not have been an opportunity to learn whether the face is congruent or incongruent on the first trial it is observed. Therefore we would not expect to see a difference between validity conditions in trial periods 2 and 3 of block 1. As such, block 1 was omitted from the analysis of trial periods 2 and 3 to retain sensitivity to effects of validity (see Figures 13 & 14 for EMG activity across blocks). These trial periods were therefore analysed separately within each muscle with a 2 × 2 × 2 × 4 mixed ANOVA, with the between-subjects factor of trust effect and within-subjects factors of validity, face gender and block (2-5). In contrast, because trial periods 4 and 5 occur during/ after the gaze-cue and target presentation, they were analysed with all 5 blocks included in a 2 × 2 × 2 × 5 mixed ANOVA. Interactions with trust effect were followed up with separate within-subjects ANOVAs in each group.

### Corrugator

In trial period 2 there was a significant main effect of validity, $F$ (1,48) = 5.478, $p$ = .023, $\eta_p^2$ = .102. This was due to the no-trust effect participants showing a larger response to congruent compared to incongruent faces, as shown by the significant interaction between trust effect and validity, $F$ (1,48) = 4.160, $p$ = .047, $\eta_p^2$ = .080. The ANOVAs conducted for each group separately showed the effect of validity was significant in the no-trust effect participants, $F$ (1,22) = 6.95, $p$ = .015, $\eta_p^2$ = .24, but was not significant in the trust effect participants, $F$ (1,26) = .064, $p$ = .802, $\eta_p^2$ = .002. In trial period 3 there was also a significant interaction between trust effect and validity, $F$ (1,48) = 4.407, $p$ = .041, $\eta_p^2$ = .084, which was again due to the no-trust effect participants showing a larger response to congruent compared to incongruent faces, $F$ (1,22) = 5.34, $p$ = .031, $\eta_p^2$ = .195, and the absence of such an effect of validity in the trust effect participants, $F$ (1,26) = .031, $p$ = .861, $\eta_p^2$ = .001. During trial period 4 there was a marginal main effect of validity, $F$ (1,48) = 3.96, $p$ =.052, $\eta_p^2$ = .076, owing to a larger response to incongruent compared to congruent gaze-cues. This was qualified by a significant interaction between trust effect and validity, $F$ (1,48) = 4.21, $p$ = .046, $\eta_p^2$ = .081. Separate ANOVAs in each group revealed a significant difference between congruent and incongruent trials in those individuals who showed a trust effect, $F$(1,26) = 6.48, $p$ =.017, $\eta_p^2$ = .200, and the absence of a significant difference in those who did not produce trust effects, $F$ (1,22) = .003, $p$ = .955, $\eta_p^2$ = .000. See Figures 11 and 12.

### Zygomaticus

In the zygomaticus there was also a significant main effect of validity in trial period 4, $F$ (1,48) = 11.0, $p$ = .002, $\eta_p^2$ = .186, again owing to a larger response to incongruent compared to congruent gaze-cues. However, there was no interaction between trust effect and validity, $F$ (1,48) = .341, $p$ = .562, $\eta_p^2$ = .186 (see Figures 13 and 14). No other main effects or interactions approached significance.

Despite several novel findings, potential technical issues need to be considered. The difference in the amplitude of the EMG responses to gaze-cues is unlikely to be explained as activity generated by concomitant eye-movements. Although previous studies have shown that gaze-cues can trigger micro saccades in the direction of the gaze, the changes in eye-position are small (Deaner & Platt, 2003). Larger, overt eye-movements are required to fixate the object beside the face and this is required for both congruent and incongruent cueing. Furthermore, the differential amplification of the bipolar electrode montage has the effect of cancelling out common signals from distant sources and the electrodes are placed away from the optimal sites used to record lateral eye-movements. Eye blinks are unlikely to explain the effects as such artefacts occur at a frequency resolution below the threshold of the high pass filter (~20 Hz). It is also unlikely that the differences are due to larger errors for incongruent compared to congruent cues. The overall differences in the number of error trials between conditions were small (<1) and were removed from all analyses. In addition, the errors appeared to be equally distributed across all 16 face identities.

### Individual Differences in Anxiety

Anxiety data were collected for 49 of the participants whilst the electrodes settled onto their skin. Using independent samples t-tests, there was no significant differences between trust effect and no-trust effect participants in terms of trait, $t\ (47) = .783$, $p = .438$, 95% CIs [-3.86, 8.87], or state anxiety, $t\ (47) = .142$, $p = .888$, 95% CIs [-4.83 5.56]. This is somewhat surprising, as it might have been predicted that individuals scoring higher on anxiety would be more sensitive to learning the contingencies between identity and gaze.

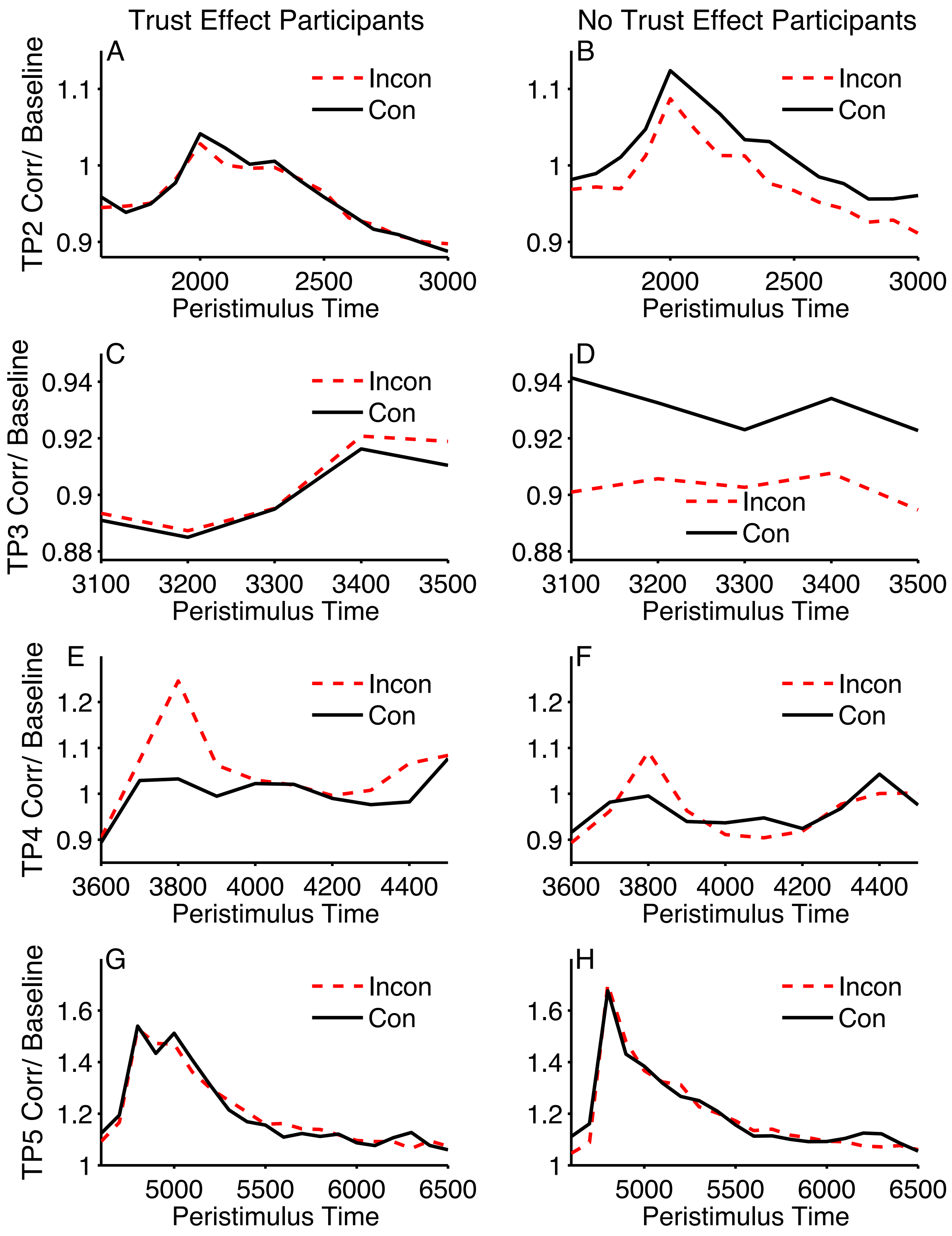


*Figure 11.* Mean stimulus-locked corrugator activity on congruent (solid line) and incongruent trials (dashed line) for trust effect (left panels) and no trust effect (right panels) participants across trial periods 2, 3, 4 and 5 (rows). EMG units on the Y-axis represent the ratio of activity to baseline (fixation).

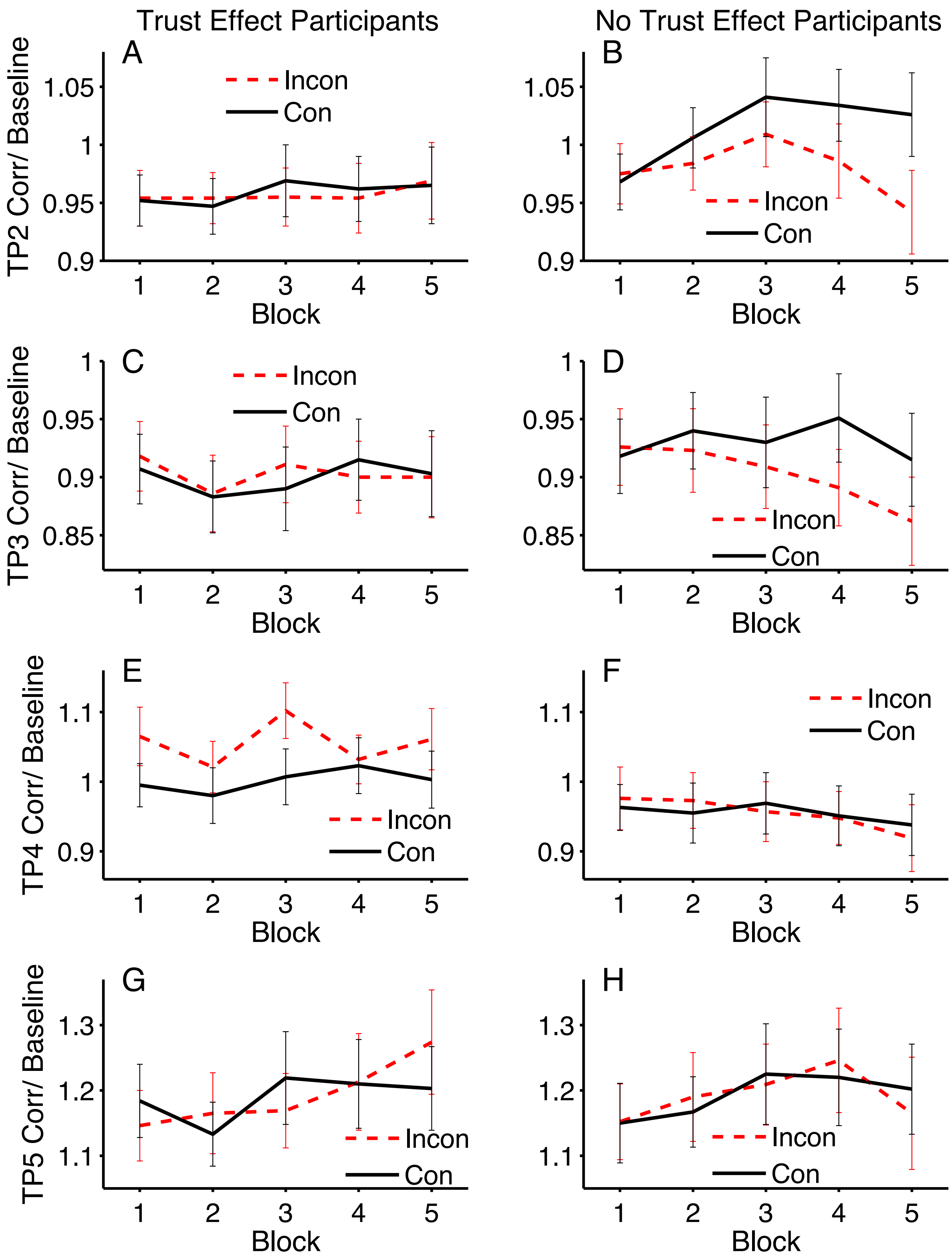


*Figure 12.* Mean corrugator activity across blocks for trust effect participants (left panels) and no-trust effect participants across trial periods (rows). Dashed lines represent incongruent and solid lines congruent. EMG units on the Y-axis represent the ratio of activity to baseline (fixation). Error bars show +/- 1 standard error of the mean.

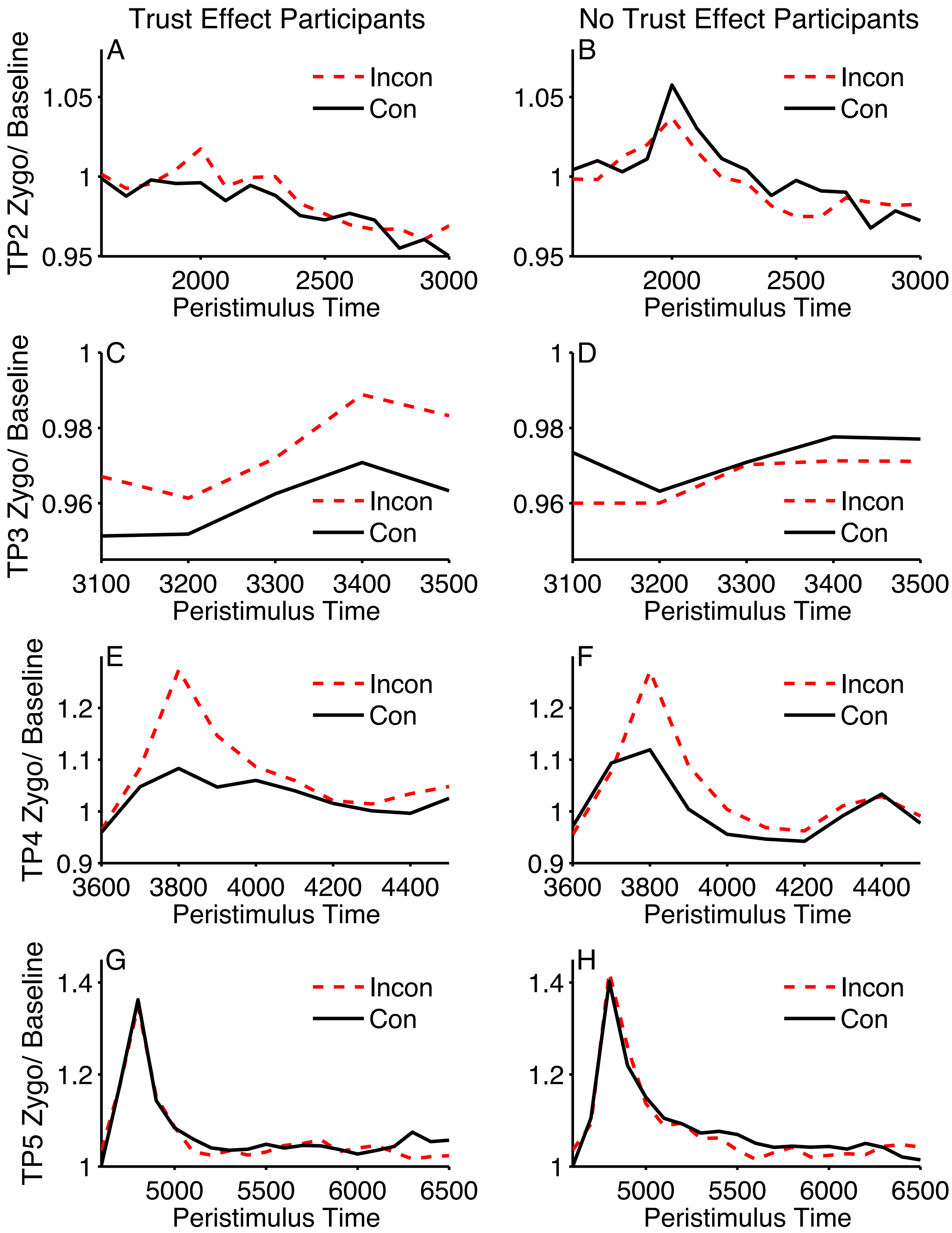


*Figure 13.* Mean stimulus-locked zygomaticus activity on congruent (solid line) and incongruent trials (dashed line) for trust effect (left panels) and no trust effect (right panels) participants across trial periods 2, 3, 4 and 5 (rows). EMG units on the Y-axis represent the ratio of activity to baseline (fixation).

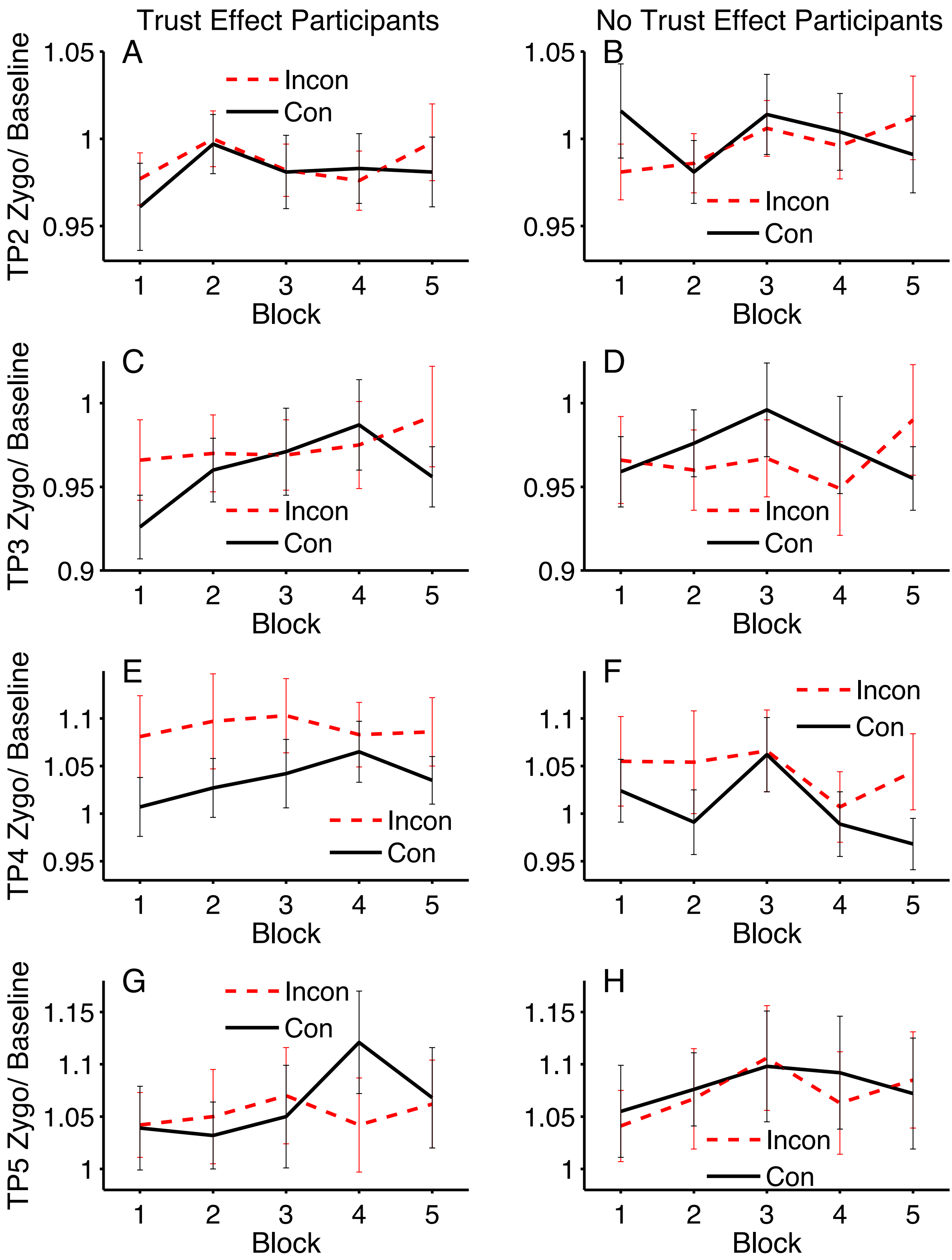


*Figure 14.* Mean zygomaticus activity across blocks for trust effect participants (left panels) and no-trust effect participants across trial periods (rows). Dashed lines represent incongruent and solid lines congruent. EMG units on the Y-axis represent the ratio of activity to baseline (fixation). Error bars show +/- 1 standard error of the mean.

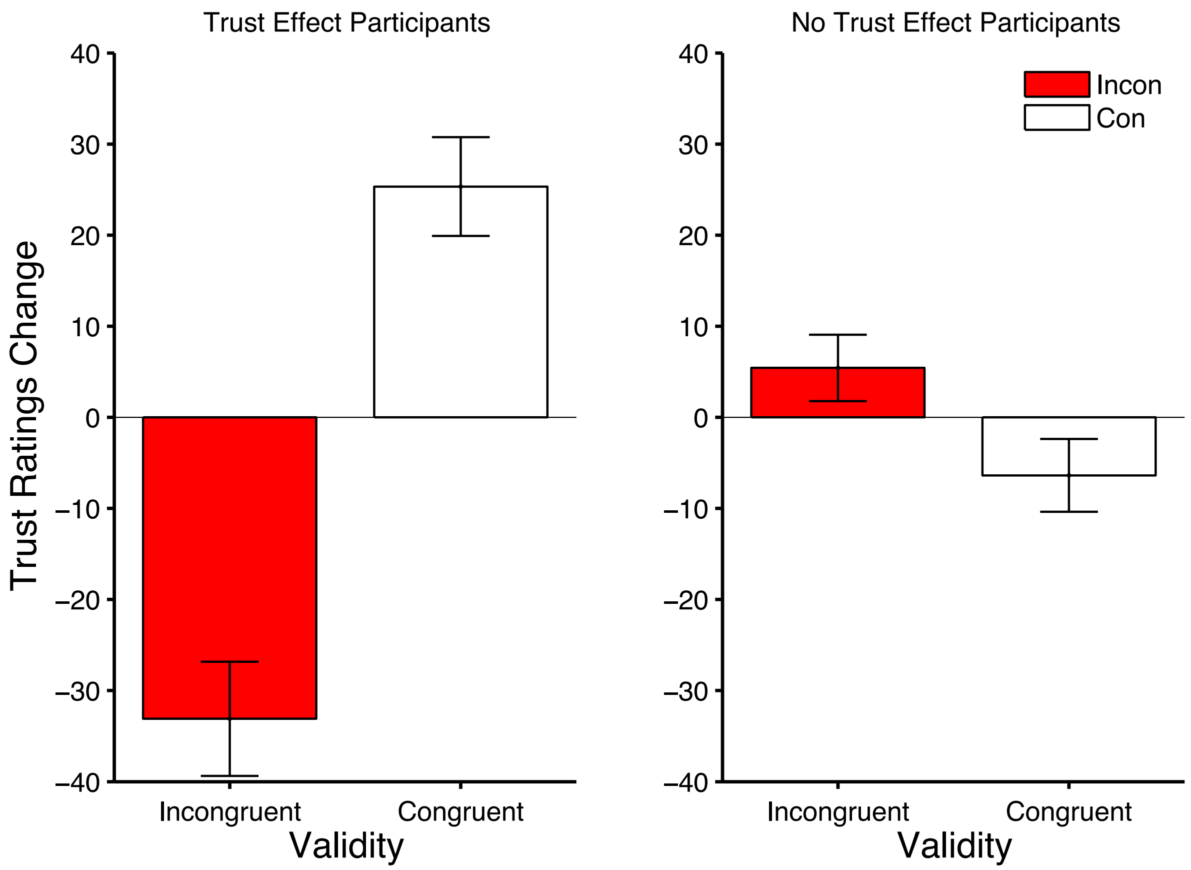


*Figure 15.* Mean change in trust ratings (end-beginning) for congruent and incongruent conditions in trust effect participants (left panels) and no trust effect participants (right panels). Error bars show +/- 1 standard error of the mean.

## Discussion

The current study is the first to use EMG to examine implicit emotional facial reactions to gaze-cues and the relation to individual differences in learning face evaluations of trustworthiness. In addition to standard gaze-cueing effects (Driver et al., 1999; Friesen & Kingstone, 1998; Frischen et al., 2007), we replicated the effects of these cues on trust judgements as incongruent faces became judged less trustworthy than congruent faces (Bayliss & Tipper, 2006; Bayliss et al., 2009). As exactly the same pattern of changes in ratings was observed in experiments 1 and 2, the contribution of the pupil manipulation to the learning of trust in experiment 1 must have been minimal because both experiments were the same apart from omission of this variable in experiment 2.

Although the changes in trust due to congruent and incongruent gaze-cueing was robust across analysis of all participants, some individuals did not change trust ratings in the expected manner. Of note is the marked similarity of gaze-cueing effects in the two groups of participants who differed in later judgements of trust. Hence, it is not simply the impairment of processing produced when a person gazes away from a target object that influences changes in trust ratings. Rather, we proposed something more is required, and that was an emotional reaction, reflected in an embodied state measured via EMG, to the congruence of gaze-cues. Participants who experienced these emotional state changes would produce changes in trust judgements. In an attempt to measure embodied emotional reactions during gaze-cueing we recorded facial EMG from the corrugator and zygomaticus muscles.

During trial period 4 when cueing was taking place as the eyes gazed either towards or away from a target object, larger activity in the corrugator and zygomaticus was observed in response to incongruent compared to congruent faces. The fact that the response was observed in both muscles and was greater to incongruent faces suggests that this activity reflects a negative emotional reaction. Previous research has shown that the corrugator is responsive to unpleasant stimuli, whereas the zygomaticus shows a bivalent profile, responding to both pleasant and unpleasant stimuli, albeit somewhat less so to the latter (Armel, Pulido, Wixted & Chiba, 2009; Lang et al., 1993; Larsen et al., 2003; Sims et al., 2012). Our findings are also consistent with this pattern. At first glance, it may appear that the change in trust ratings are somewhat incompatible with the EMG in that both a positive change to congruent and a negative change to incongruent was observed. In contrast, EMG activity was larger for incongruent faces in both the corrugator and zygomaticus during trial period 4, which raises the question as to how the positive change to congruent faces is elicited if the zygomaticus effect primarily reflects a negative reaction to incongruent cueing.

We argue that the positive trust values associated with congruent faces observed in both experiment 1 and 2 are due to contrast effects with the incongruent faces. First, the greater EMG activity to incongruent faces is similar to experiment 1 where the LPP was larger to incongruent faces as blocks progressed. Again, this suggests that deceptive behaviour is given greater weight than trustworthy behaviour. Second, the increased salience of the negative incongruent faces produces a contrast effect during the later overt/ explicit judgements of trust. That is, there are comparison processes in memory, where the ratings to congruent faces are made in reference to incongruent faces. Such contrast effects have been observed by others. For example, adaptation to very trustworthy and very untrustworthy faces was found to bias perception of less extreme test faces, such that they became judged less and more trustworthy, respectively (Wincenciak, Dzhelyova, Perrett & Barraclough, 2013). There is also evidence that the EMG can reflect affective contrast. Larsen and Norris (2003) have shown that corrugator EMG activity is larger to moderately pleasant pictures when presented among extremely pleasant pictures compared to mildly pleasant pictures.

However, the main focus of this experiment was the identification of what differentiated those people whose later judgements of trust were determined by the validity of the faces. As noted, the standard cueing effects were similar in participants who did and did not show changes in trust ratings. However, recording of embodied emotional reactions during gaze-cueing did relate to the individual differences in changes in trust ratings. In trial period 4, trust effect participants showed a larger EMG reaction to incongruent compared to congruent gaze-cues than participants who did not show trust effects. In addition, in trial periods 2 and 3, the no trust effect participants showed larger corrugator activity to congruent compared to incongruent than trust effect participants. This is consistent with the ratings of the no trust participants who showed a trust effect in the opposite direction to the trust participants as evidenced by a significant time × validity interaction ($F\,(1, 22) = 5.98$, $p = .023$, $\eta_p^2 = .214$), where trust of incongruent faces increased and trust of congruent faces declined. This reverse effect was weaker than the standard effect (see Figure 15).

Therefore, the reverse effect in the no trust participants is not inconsistent with our hypotheses and may well be due to the procedure of dividing participants into trust and no trust groups. It shows that EMG reactions are related to evaluative ratings and that selecting participants based on ratings can create corresponding changes in EMG. As to why the effects for the trust and no trust participants occurred in separate trial periods is unclear. As the trust effect participants are reacting at the time of the gaze-cue, the emotional signal may be based more accurately on the congruency of the cue. In contrast, the reaction in trial periods 2 and 3 is not as clearly based on the congruency of the gaze-cue, as no target is present. This may be a mistrust of congruent faces that is not based on gaze-cues and is large in magnitude due to the effect of selecting and summing together individuals who are biased in this way. An alternative explanation could be that the no trust effect participants have personality types that make them prone to behaving deceptively themselves. They may accept deception as a moral norm, making them less sensitive to untrustworthy behaviour. The larger activity to congruent gaze-cueing may be because the participant's tendency to behave untrustworthily creates conflict or guilt when the other person is trustworthy.

In conclusion, here we provide initial evidence from facial EMG for the role of embodied emotion states/ somatic markers and individual differences in learning face evaluations of trustworthiness from gaze-cueing. The observed effects were driven by a negative emotional reaction during the observation and response to incongruent faces within individuals who showed evidence of learning to mistrust these faces more than congruent faces.

# Chapter 4

## Experiment 3: Effects of emotion induction via priming of social exclusion and inclusion on learning of trust from eye-gaze cues

Experiments 1 and 2 have shown effects of gaze-cues on implicit measures of emotion and a relation to the learning of trustworthiness. Experiment 3 approaches the core question of whether emotion mediates learning of trust from gaze using a slightly different approach. Instead of measuring emotion in response to identity-contingent gaze, experiment 3 examines whether inducing participants to feel more or less positive or negative influences learning of trust from gaze-cues given the role for emotion already established in experiments 1 and 2. Indeed, Dunn and Schweitzer (2005) found that participants who were induced to feel angry or sad trusted others less than when they were induced to feel happy or grateful although it is unclear whether such emotions influence learning of trust. We were interested in the effects of social emotions which derive from the experienced, recalled or anticipated representation of other peoples thoughts, feelings or actions (Hareli & Parkinson, 2008).

To achieve this, a mood induction procedure that involved priming participants with the concept of social inclusion or exclusion was used. Social exclusion can be experienced as extremely aversive. One way in which social inclusion and exclusion has been studied is through use of the cyber ball paradigm- a virtual ball tossing game between the participant and two other avatars the participants believe are controlled by other people. When participants are excluded, the other players rarely pass them the ball, whereas when they are included they are passed the ball more frequently (Williams, Cheung & Choi, 2000). When participants are excluded in ways similar to this, they report a decrease in positive mood ratings and an increase in anger ratings whereas the positive changes to inclusion are less extreme (Blackhart et al., 2009; Seidel et al., 2013). In fMRI studies, social rejection is paralleled by activation in the ACC, which is believed to reflect the affective qualities pain (Rotge et al., 2014). We predict that the priming of exclusion will sensitise participants to the positive and negative connotations of identity-contingent gaze-cues producing an increased trust effect and EMG activity differences compared to inclusion.

In experiment 3, separate groups of participants were primed with either inclusion or exclusion in three ways. Firstly, before cueing, participants were presented with a cartoon clip depicting a group of shapes playing together whilst another shape was either excluded or not excluded from their activities (Over & Carpenter, 2009). Such shapes are typically perceived as having agency and participants can use their motion to infer mental states (Castelli et al., 2000). Secondly, before cueing, but after the cartoon, participants were asked to introspect upon a time in which they felt included or excluded from a groups activity. Thirdly, in between blocks, participants observed images of social scenes in which all of the individuals were included in a group's activity or in which one individual was excluded from the activities.

## Methods

### Participants

There were 62 participants in total, half of which were designated to the inclusion group and half to the exclusion group. The inclusion group had a mean age of 20 (*SD*=3) and there were 23 females and 27 right-handers. The exclusion group had a mean age of 19 (*SD*=1) and was made up of 23 females and 26 right-handers. All participants were neurologically normal and had normal or corrected-to-normal vision. The ethics committee of Bangor University approved all procedures.

### Stimuli and Apparatus

The face and target object stimuli were the same as that used in experiment 2. Participants observed one of two animated cartoons before the cueing phase began. The two video clips were of cartoon shapes in either an inclusion or exclusion scenario (Over & Carpenter, 2009). In the exclusion scenario, three shapes were depicted playing in a group together whilst excluding a fourth shape from joining their circle. The video also depicted two shapes playing football together whilst leaving out a third. The inclusion cartoon was exactly the same except that there was no excluded shape. Instead, a fly took its place and flew randomly about the play area (see Figure 16). Both the inclusion and exclusion cartoons lasted 1 minute and 55 seconds. In between blocks participants were presented with images of social situations, which, in the exclusion condition, depicted people being excluded from a group's activity. In the inclusion condition all the individuals were engaged in a group activity. There were 4 different images of each type (see Figures 17 & 18). Each image had a width of 700 pixels but the height varied from 325 to 593 pixels. The cueing faces were the same as those used in experiment 1 and 2 and were 504 × 600 pixels in size. The target objects were 350 × 263 pixels in size. The experiment was run with a screen resolution at 1920 × 1080 on a 24" Samsung SyncMaster BX2431 LED display, which was 569 × 342 mm in dimensions and had a 500 Hz refresh rate. The experiment was programmed and run in E-Prime 2.0 to enable video playback. EMG was collected continuously from the corrugator and zygomaticus muscles and processed the same way as in experiment 2.

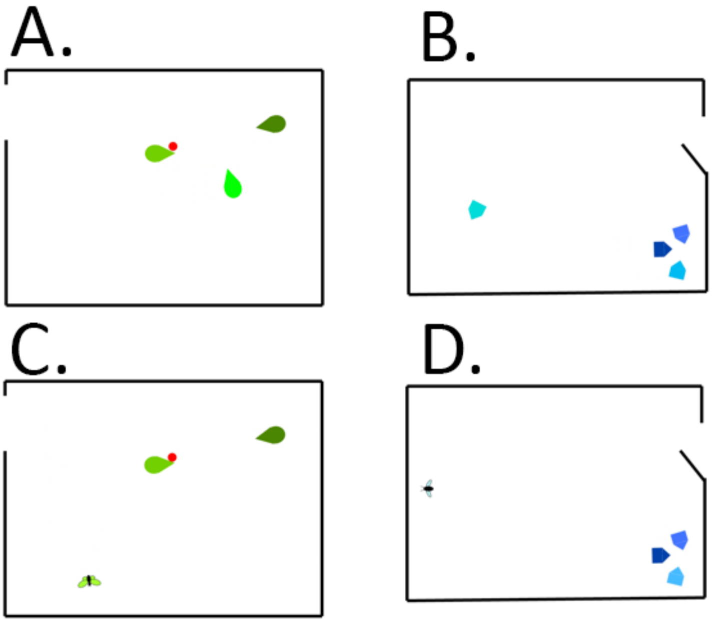


*Figure 16*. Still images from the exclusion (A and B) and inclusion (C and D) cartoons.

*Figure 17*. Example of one of the inclusion images presented in experiment 3.

*Figure 18.* Example of one of the exclusion images presented in experiment 3.

**Design and Procedure**

The procedure for this experiment was exactly the same as that of experiment 2 with only three variations. Firstly, before beginning the first block of the experiment, participants watched either the inclusion or exclusion cartoon. Secondly, they were then given 30 seconds to close their eyes and recall a time in which they felt included or excluded from a groups activities. Specifically, participants were requested to: *"Recall a time when you were disappointed that your friends left you out of their activities"* or *"Recall a time when you were pleased that your friends included you in their activities"*. After 30 seconds had elapsed, participants heard a bell to signify that they could open their eyes and proceed with the experiment. If they could not remember within 30 seconds they were given an extra 30 seconds if needed. As a manipulation check, participants were then asked if they successfully recalled a time they felt included or excluded. All participants reported that they successfully managed to retrieve and reflect upon a memory of being included or excluded. Thirdly, at the start of each block, apart from the first, participants viewed images showing inclusion or exclusion. The images were presented for 30 seconds each. Before the experiment began, and while the electrodes settled onto the face, participants completed the autism quotient (Baron-Cohen et al., 2001) and empathy quotient (Baron-Cohen & Wheelwright, 2004).

## Data Screening Protocol and Analysis

Table 7 shows the mean percentages of trials removed from analyses. All trials on which participants made an error or did not respond (*M* = 2.89%, *SD* = 2.71% of trials) were removed from reaction time and EMG analyses along with trials with reaction times above or below two standard deviations from each participant's mean (*M* = 4.53%, *SD* = 1.16% of trials). In addition, any remaining reaction times larger than 1500 ms were removed (*M* = 3.9%, *SD* = 6.84% of trials). The first trial of the experiment and trials immediately succeeding error trials were removed from the EMG data to eliminate facial reactions to novelty and errors (*M* = 3.35%, *SD* = 2.4% of trials). Trials with artefacts observed were removed from the corrugator (*M* = 8.2% of trials, *SD* = 2.91%) and zygomaticus (*M* = 12.1% of trials, *SD* = 3.4%) analyses separately. Paired samples t-tests were used to test for differences in the number of errors, outliers and artefacts between incongruent and congruent conditions. There were no significant differences between conditions in terms of the number of errors, *t* (61) = -1.8, *p* = .08, 95% CIs [-.62, .04], or zygomaticus artefacts, *t* (61) = 1.3, *p* = .195, 95% CIs [-.24, 1.2]. However, there were significant differences in the numbers of outliers above or below two standard deviations of the participants mean, *t* (61) = 2.36, *p* = .021, 95% CIs [.07, .9], outliers above 1500 ms, *t* (61) = 2.2, *p* = .031, 95% CIs [.04, .73], and corrugator artefacts, *t* (61) = 2.27, *p* = .027, 95% CIs [.08, 1.2]. Overall, there was no significant difference in the number of trials between conditions in the zygomaticus, *t* (61) =-1.12, *p* = .265, 95% CIs [-1.46, .41], and corrugator analyses, *t* (61) =-.74, *p* = .462, 95% CIs [-1.2, .55]. All participants were divided into those who showed trust effects and those who did not. This was entered into the reaction time and EMG ANOVAs as a between-subjects factor. Overall, 31 participants showed trust effects whereas 31 participants did not.

Table 7. *Means and standard deviations of the percentages of errors, outliers and artefacts removed across conditions.*

| | Errors | Outliers | Corrugator Artefacts | Zygomaticus Artefacts |
|---|---|---|---|---|
| Congruent | 1.3(1.5) | 3.8(3.39) | 4.4(1.93) | 6.3(2.4) |
| Incongruent | 1.6(1.5) | 4.7(3.84) | 3.8(1.74) | 5.9(2.03) |

## Results

### Evaluations of Trustworthiness

The trust ratings were analysed with a 2 × 2 × 2 × 2 mixed ANOVA with the between subjects factor of inclusion/ exclusion group and within subjects factors of time, validity and face gender. There was a significant main effect of validity $F$ (1, 60) = 20.9, $p$ <.0001, $\eta_p^2$ = .258, which was qualified by a significant time × validity interaction $F$ (1, 60) = 10.8, $p$ = .002, $\eta_p^2$ = .152. This was due to trust ratings being more negative to incongruent compared to congruent after, relative to before, cueing. There was also a significant effect of face gender, $F$ (1, 60) = 7.69, $p$ = .007, $\eta_p^2$ = .114, owing to the fact that females were rated as more trustworthy than males. This effect was larger at the beginning of the experiment compared to the end as shown by the significant time × face gender interaction, $F$ (1, 60) = 5.1, $p$ = .027, $\eta_p^2$ = .079. However, there were no main effects or interactions with inclusion/ exclusion group. There was no significant time × validity × inclusion/ exclusion group interaction $F$ (1, 60) = .88, $p$ = .353, $\eta_p^2$ = .014, although the trust effect evident in the change scores does appear to be larger in the exclusion participants, driven particularly by the incongruent condition (see Figure 19).

As in previous experiments, separate, 2 × 2 × 2 ANOVAs with within-subjects factors of validity and face gender and the between-subjects factor of inclusion/ exclusion group were run on the beginning and end ratings to formally identify the sources of the main effects and interactions described above. At the beginning, there were no significant effects apart from face gender, $F$ (1, 60) = 12.58, $p$ = .001, $\eta_p^2$ = .173, whereas at the end, there was only a significant effect of validity, $F$ (1, 60) = 19.24, $p$ < .0001, $\eta_p^2$ = .24.

### Gaze-Cueing Reaction Times

The reaction times were submitted to a 2 × 2 × 2 × 5 mixed ANOVA with inclusion/ exclusion group as the between-subjects factor and face gender, validity and block as the within-subjects factors. There was a significant main effect of validity $F$ (1, 60) = 36.7, $p$ < .0001, $\eta_p^2$ = .38, owing to slower reaction times on incongruent trials (*M* = 861.879, *SEM* = 15.6) compared to congruent trials (*M* = 831.106, *SEM* = 15.6) (see Figure 20, all panels). There was also a significant effect of block, $F$ (1, 60) = 72.9, $p$ < .0001, $\eta_p^2$ = .549. This effect of block was also significant as a linear trend $F$ (1, 60) = 129.12, $p$ < .0001, $\eta_p^2$ = .683, showing reaction times generally increased across blocks. However, there was no significant interaction of validity with inclusion/ exclusion $F$ (1, 60) = .39, $p$ = .534, $\eta_p^2$ = .006, or a validity × block × inclusion/exclusion interaction $F$ (4, 240) = 1.6, $p = .188$, $\eta_p^2 = .025$ (see Figure 20, top and bottom panels). No other effects of reached significance. Trust effect was entered into a separate ANOVA as a between-subjects factor. There were no further main effects or interactions with trust effect (see Figure 20, left and right panels).

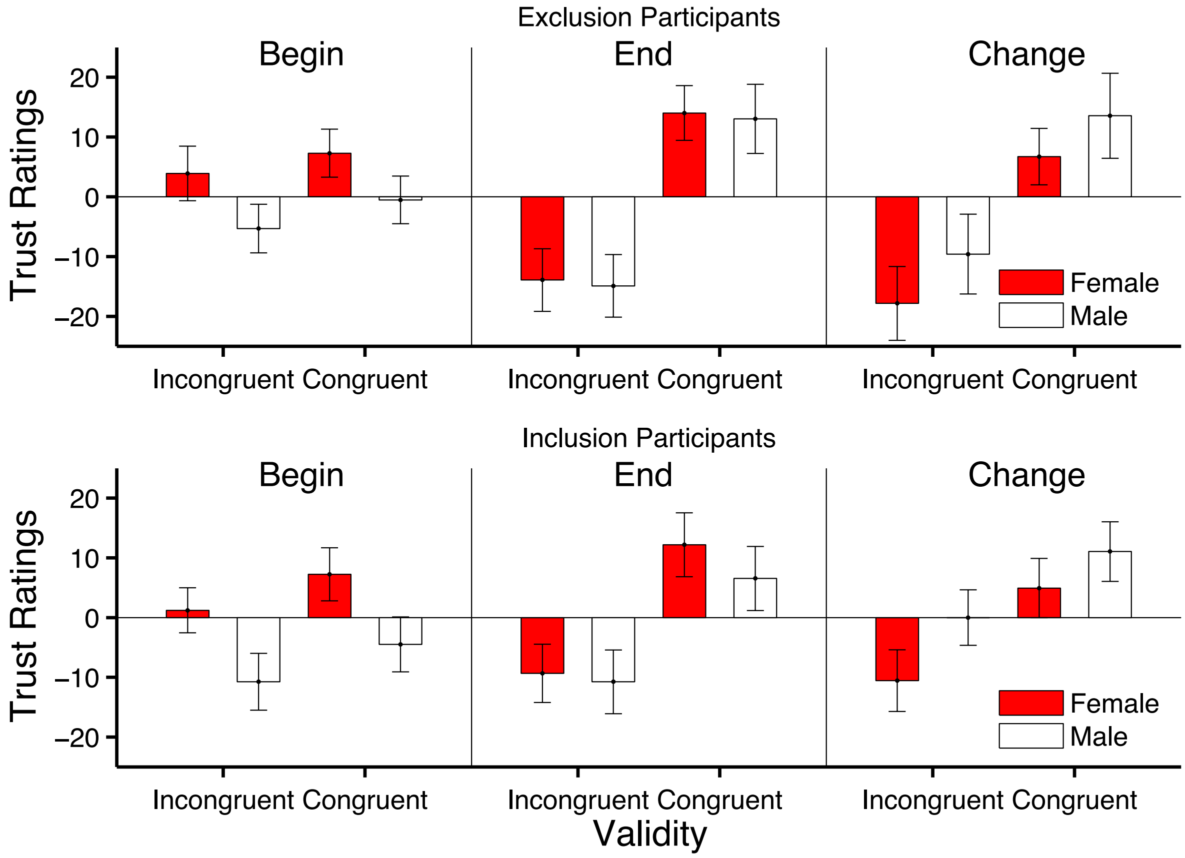


*Figure 19.* Trustworthiness ratings before (left panels) and after (middle panels) cueing and the change in ratings from before to after cueing (end-beginning) (right panels) for exclusion participants (top panels) and inclusion participants (bottom panels) in experiment 3. Error bars show +/- 1 standard error of the mean.

**Facial Electromyography (EMG)**

The EMG data was analysed in the same manner as in experiment 2. Trial periods 2 and 3 were analysed using a 2 × 2 × 2 × 4 mixed ANOVA with the between-subjects factor of inclusion/ exclusion group and the within-subjects factors of validity, face gender and block. The block factor excluded the first block as trial periods 2 and 3 occur before the gaze-cue and so participants will not have acquired any knowledge of the contingency between faces and gaze-cues in these periods. In contrast, trial periods 4 and 5 were analysed with a 2 × 2 × 2 × 5 mixed ANOVA that included all 5 blocks as validity information is acquired in these trial periods. These analyses showed that there was a significant effect of validity in both the corrugator, $F$ (1, 60) = 11.6, $p$ = .001, $\eta_p^2$ = .162, and zygomaticus, $F$ (1, 60) = 11.04, $p$ = .002, $\eta_p^2$ = .155, in trial period 4 due to larger responses to incongruent compared to congruent (see Figures 21, 22, 23 and 24, panels E and F). No other main effects or interactions approached significance. All analyses were reproduced with trust effect as the between-subjects factor. However, there were no further significant interactions with trust effect.

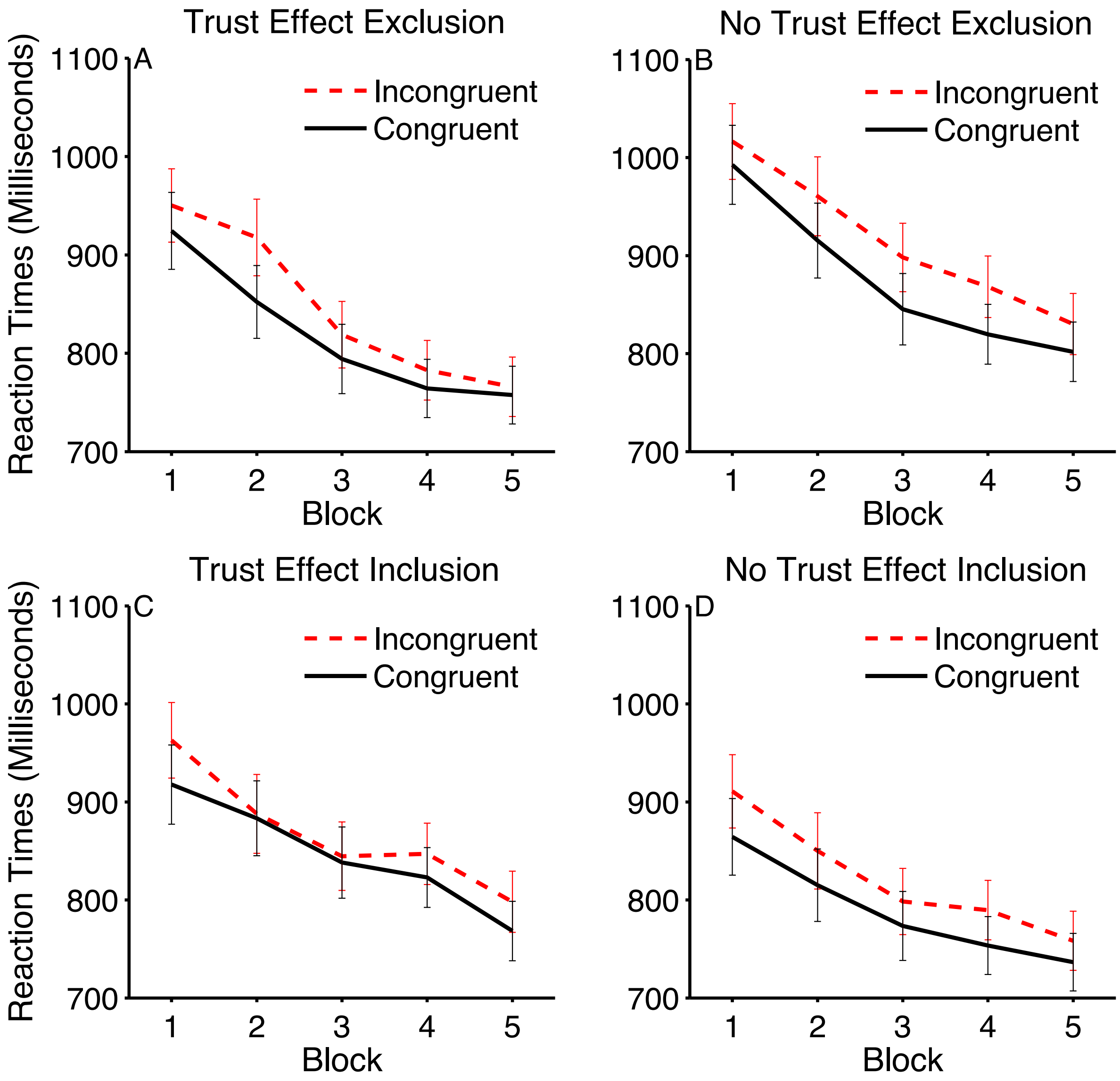


*Figure 20*. Gaze-cueing reaction times across blocks shown separately for trust effect participants (left panels) and no trust effect participants (right panels) in the exclusion (top panels) and inclusion groups (bottom panels) in experiment 3. Error bars show +/- 1 standard error of the mean.

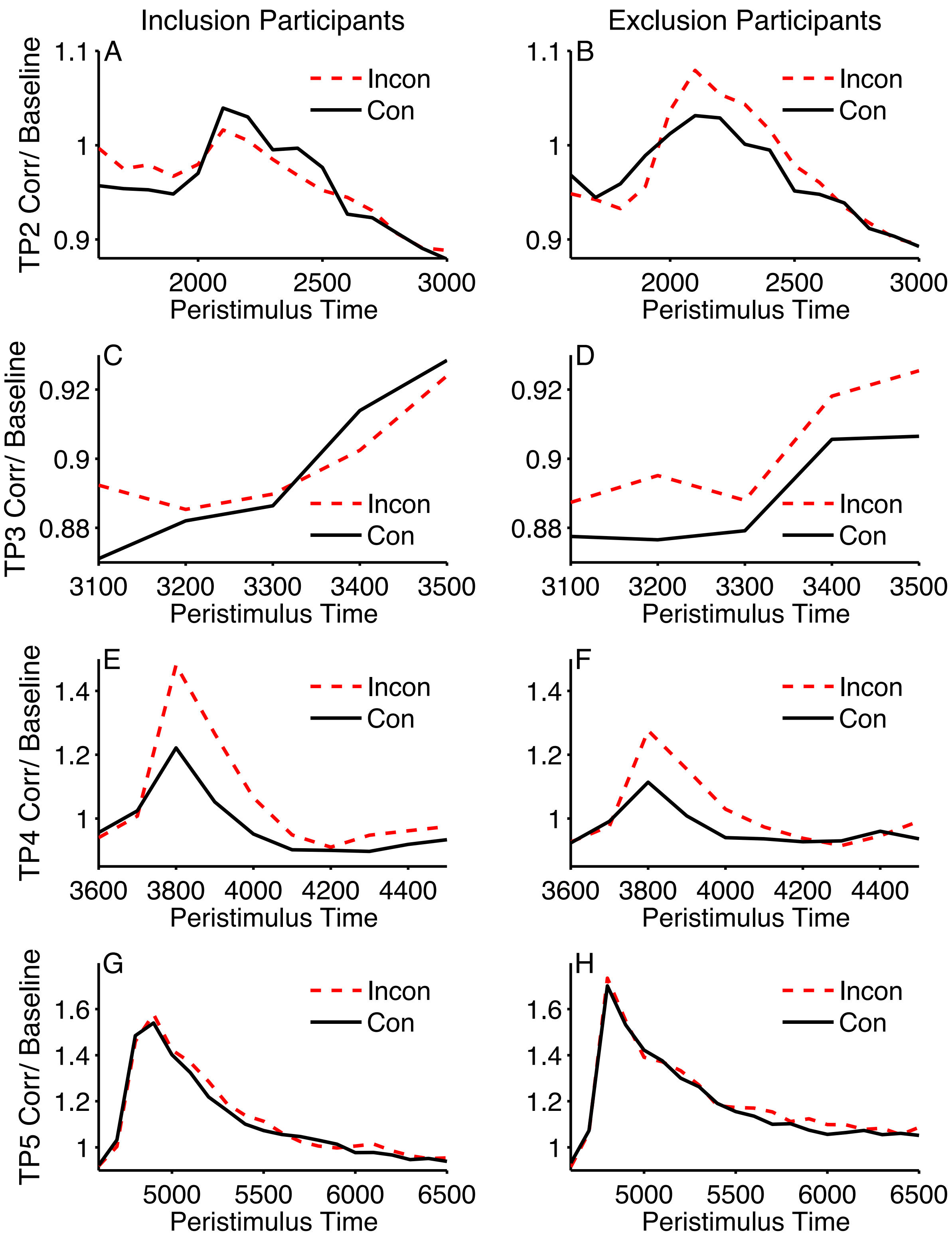


*Figure 21.* Mean stimulus-locked corrugator activity on congruent (solid line) and incongruent trials (dashed line) for inclusion participants (left panels) and no trust effect participants (right panels) across trial periods 2, 3, 4 and 5 (rows). EMG units on the Y-axis represent the ratio of activity relative to baseline (fixation).

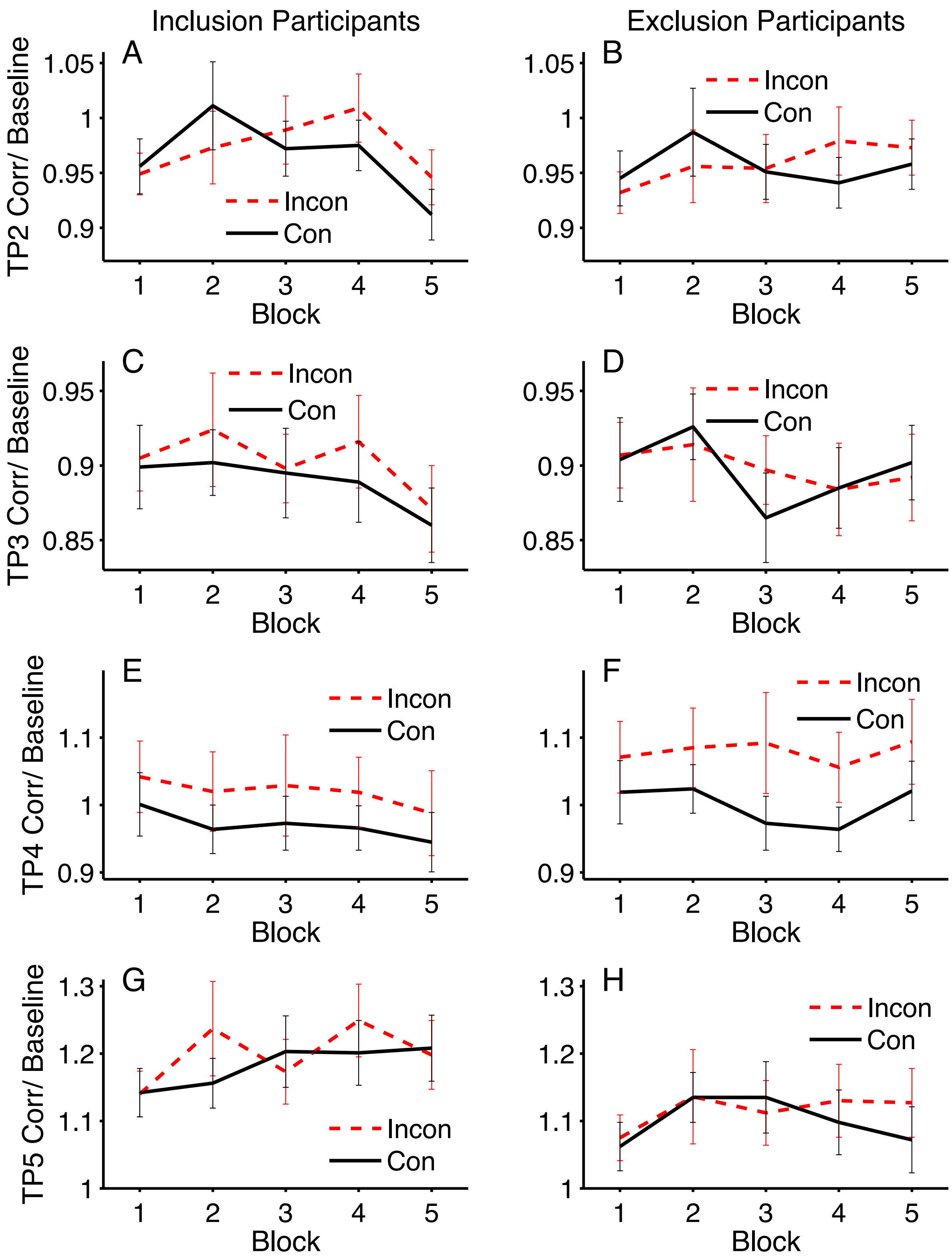


*Figure 22.* Mean corrugator activity across blocks for inclusion participants (left panels) and exclusion participants (right panels) in trial periods 2, 3, 4 and 5 (rows). Dashed lines represent incongruent and solid lines represent congruent. EMG units on the Y-axis represent the ratio of activity relative to baseline (fixation). Error bars show +/- 1 standard error of the mean.

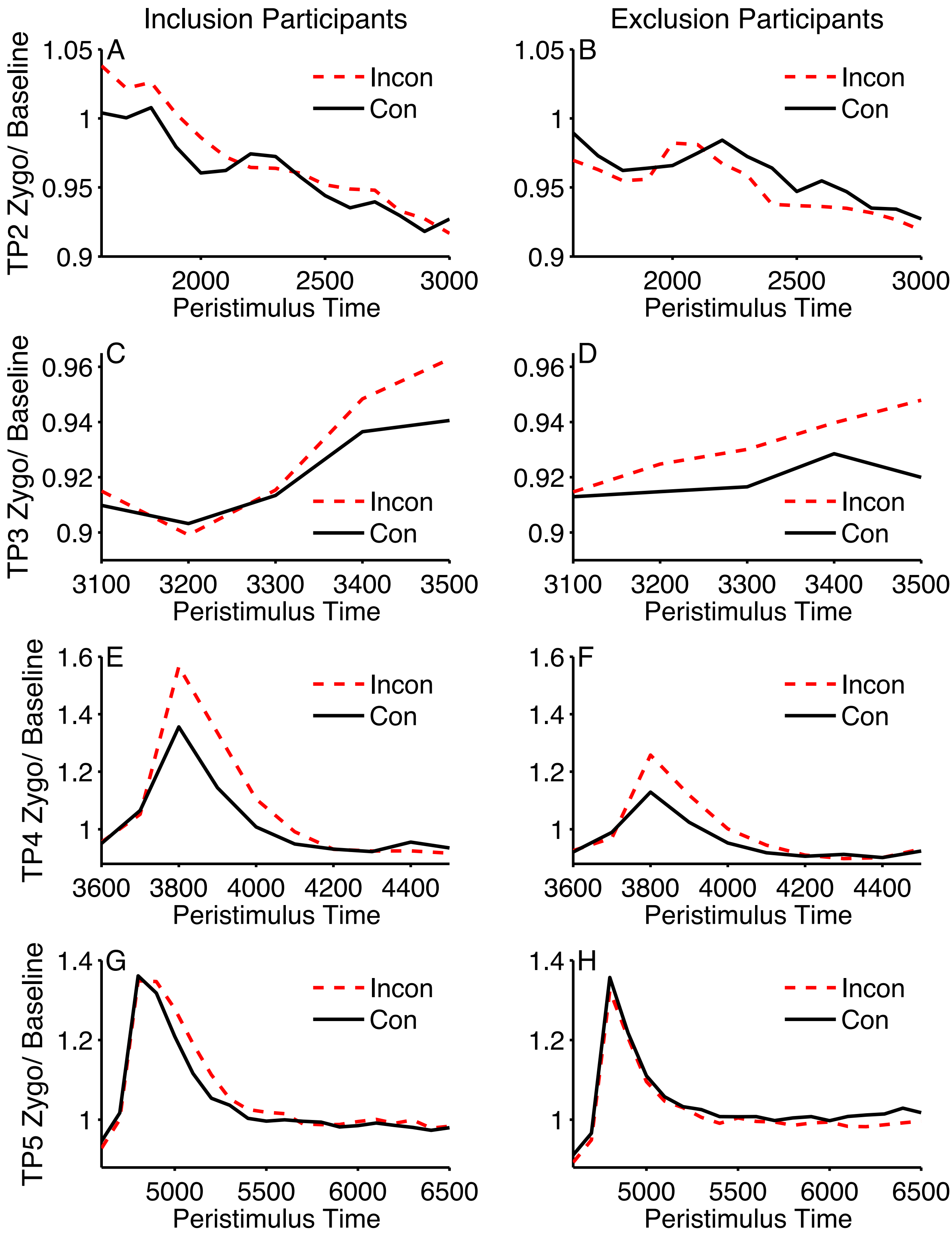


*Figure 23.* Mean stimulus-locked zygomaticus activity on congruent (solid line) and incongruent trials (dashed line) for inclusion participants (left panels) and exclusion participants (right panels) across trial periods 2, 3, 4 and 5 (rows). EMG units on the Y-axis represent the ratio of activity relative to baseline (fixation).

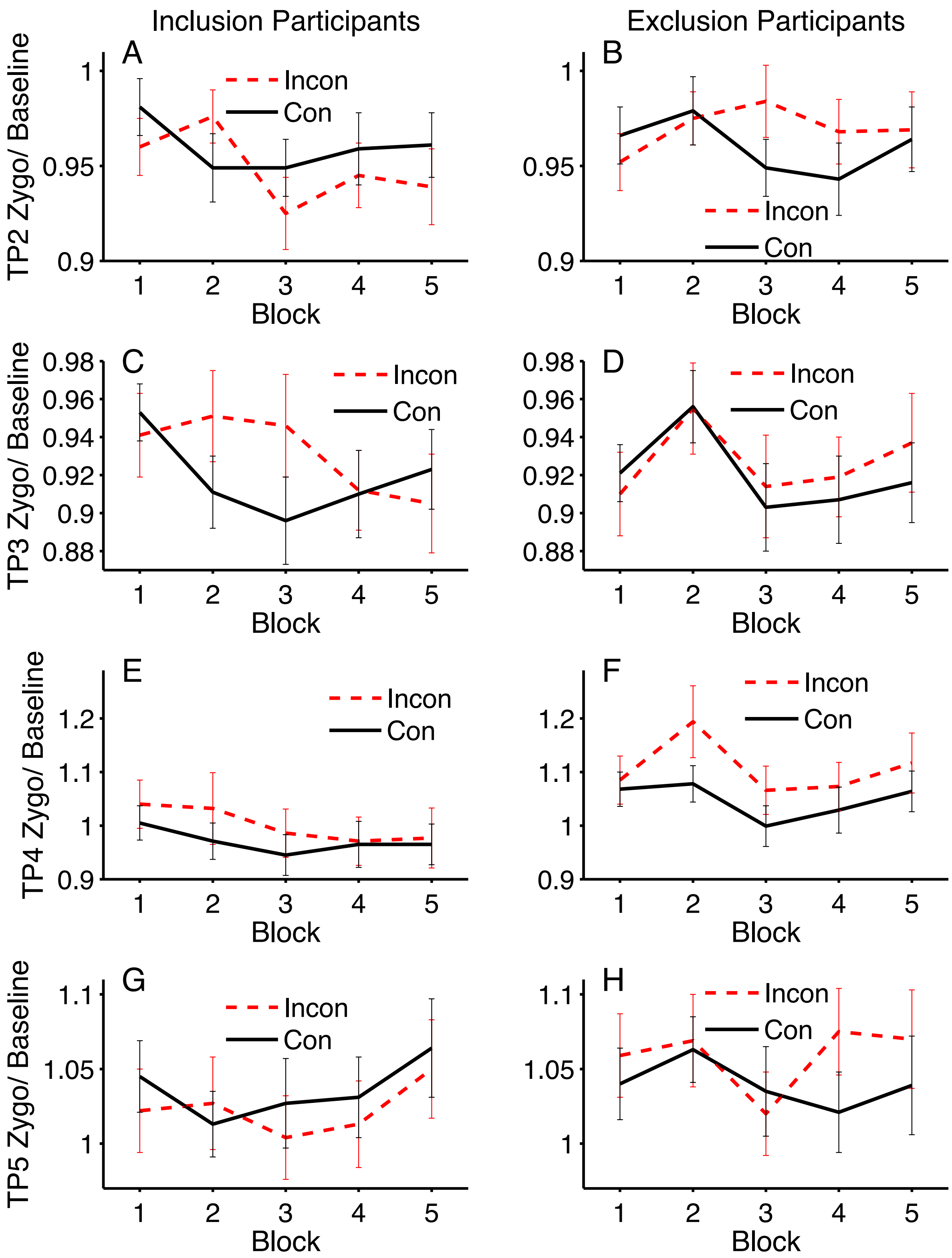


*Figure 24.* Mean zygomaticus activity across blocks for inclusion participants (left panels) and exclusion participants (right panels) in trial periods 2, 3, 4 and 5 (rows). Dashed lines represent incongruent and solid lines represent congruent. EMG units on the Y-axis represent the ratio of activity relative to baseline (fixation). Error bars show +/- 1 standard error of the mean.

## Discussion

Experiment 3 investigated the effects of priming the emotions associated with exclusion and inclusion on emotional reactions to gaze cues and its relation to learning of trust. The effects of identity-contingent gaze-cues on reaction times and trust judgements were remarkably similar to experiments 1 and 2. Participants were slower to respond to incongruent faces and such faces became judged less trustworthy than congruent faces after cueing, particularly female faces. We also replicated the main effect of validity in both the corrugator and zygomaticus muscle in trial period 4, where EMG activity was larger to incongruent faces. This again confirms that gaze-cues can elicit an embodied emotional response. However, there were no effects of inclusion and exclusion or interactions with validity in both the ratings and EMG activity. This is unexpected given that exclusion is a potent elicitor of negative emotional states (Seidel et al., 2013) and that negative emotional states can reduce trustworthiness judgements of others (Dunn & Schweitzer, 2005), although it is notable that the effect was trending in the right direction. Exclusion participants tended to rate incongruent faces less trustworthy than inclusion participants.

This could be for several reasons. It has been shown that the effect of emotion induction on trust judgements is only evident when participants are not aware that their emotional states are elicited by the induction procedure (Dunn & Schweitzer, 2005). That is, when participants are not aware of the source of their emotions, they misattribute them to other sources. In the case of the current experiment, participants may have been aware that their emotional state was influenced by the exclusion/ inclusion manipulation. As such, it may not have been attributed to the gaze-cues. The lack of an interaction between inclusion/ exclusion and trust learning may actually be revealing important information about the nature of emotion/ social emotions. There may be distinct emotions for trust and rejection, consistent with the many distinct social emotions that have been proposed (Hareli & Parkinson, 2008). As such, one may not expect to see interplay between the two.

It may also be necessary to distinguish between emotion and mood. Emotions are states elicited by specific stimuli and typically decrease more rapidly after the stimulus is terminated. In contrast, mood is a much more sustained affective state. The procedure used in this experiment to alter participant's emotional state was more similar to mood induction as there were large intervals between the inclusion/ exclusion stimuli. Therefore, inclusion and exclusion could influence emotion and learning of trust from gaze-cues if the inclusion/ exclusion stimuli were presented immediately before the gaze-cueing trials so that the emotional state continues into these trials. The emotional state in trials immediately after presentation of the inclusion/ exclusion stimulus would be expected to be larger than after several of the gaze-cueing trials as the emotion dissipates. Indeed, the LPP ERP to emotional stimuli often remains sustained for several seconds after stimulus offset (Hajcak et al., 2010). Although it is known that social rejection is experienced negatively, it would be worthwhile if similar future experiments measure mood and emotion in response to the inclusion/ exclusion stimuli, which, when controlled for, may reveal effects.

In comparison to experiment 2, there were also no interactions with trust effect. This may highlight potential boundary conditions for when individual differences in emotion-related changes in trust from gaze-cues may be expressed. The interactions between individual differences in trust effects and validity observed in the EMG in experiment 2 only just passed the threshold for significance. The introduction of the social rejection manipulation may have interrupted this relationship between trust effect and EMG activity. However, as in experiments 1 and 2, learning of mistrust from incongruent gaze was greater for female faces and learning of trust from congruent faces was greater for male faces as shown. This suggests that learning of trust is modulated by participants initial expectations of how trustworthy the person should behave based on visual appearance. This consistent observation suggests that learning of trust from gaze-cues may potentially conform to prediction-error learning rules and this therefore warrants further investigation.

In summary, experiment 3 has examined whether increasing emotion/ mood states through priming exclusion or inclusion modulates the learning of trust from gaze-cues. While the overall findings are consistent with the previous experiments we did not find any effects of the exclusion manipulation. There may be several reasons for this and this does not necessarily rule out the notion that emotion mediates the learning of trust from gaze-cues.

# Chapter 5

## Experiment 4: The interaction of initial trust judgements based on facial appearance with the learning of trust from identity-contingent gaze-cues

In experiments 1, 2 and 3, the learning of trust from gaze-cues was modulated by face gender in a very similar manner. Females became judged less trustworthy than males when incongruent and males became judged more trustworthy than females when congruent. This is consistent with Bayliss et al. (2009) which showed that the effects of gaze-cues on learning trust are greater when the cueing faces display a happy expression compared to when expressing anger or neutral. These findings suggest that incongruent gaze-cues trigger more negative emotional reactions and this increases in magnitude when the initial expectancy of a positive, trustworthy, interaction is violated in a manner akin to the Rescorla and Wagner (1972) learning rule. Indeed, faces that appear more masculine and angry are judged less trustworthy than faces which are more feminine and happy (Oosterhof & Todorov, 2008). Happy faces are also more likely to be judged as gazing towards the observer when gaze is slightly deviated, especially for female faces, suggesting people are more likely to anticipate pro-social gaze behaviours (Lobmaier, Tiddeman & Perrett, 2008; Slepian, Weisbuch, Adams & Ambady, 2011). However, it is still unclear as to how appearance and gaze-cueing behaviour interacts and contributes to trustworthiness judgements.

It is known that both trustworthy appearance and behaviour can elicit neural and psychophysiological activity associated with emotion (Chapman et al., 2009; Engell et al., 2007; King-Casas et al., 2005; Marzi et al., 2014; Sanfey et al., 2003; Van't Wout et al., 2006; Winston et al., 2002; Yang et al., 2011). There have also been several experiments that have investigated the interaction between appearance-based trust and the learning of trust from behavioural information. In contrast to gaze-cueing experiments, these studies have almost unequivocally shown that appearance based cues to trust do not interact with the learning of trust from behaviour in the manner predicted. No interactions in the expected direction have been observed between faces that are high/ low in appearance based trustworthiness and the learning of trust from associated positive and negative trait words (Rudoy & Paller, 2009), behavioural descriptions (Baron et al., 2011), angry and happy facial expressions (Aguado, Román, Fernández-Cahill, Diéguez-Risco & Romero-Fereiro, 2011) or monetary gains/ losses (Chang, Doll, van't Wout, Frank & Sanfey, 2010; Kim, Choi & Jang, 2012). This could be because appearance and behaviour based trustworthiness judgements are independent and used to guide behaviour in different ways. Judgements of trustworthiness based on facial appearance have not unequivocally been shown to be predictive of trustworthy behaviour (Rule et al., 2013). This could be because if trustworthiness were diagnostic from facial appearance it would be easily exploited. Thus, it has been proposed that the best way to diagnose trust is through learning (DeSteno, 2014).

Experiment 4 aimed to provide a stronger test of the idea that the learning of trust from gaze-cues is greater when there is a mismatch between appearance-based expectations of trustworthiness and the trustworthiness of the observed gaze-cues. Instead of using faces that differed in terms of gender or expressions, we used two groups of faces that were all female and neutral but more highly polarised in terms of initial trust judgements based on facial appearance. It was expected that if the initial trustworthiness of the face interacts with the learning of trust from gaze-cues, then there should be a larger positive change in trust ratings for low trust congruent faces compared to high trust congruent faces and a larger negative change in ratings for high trust incongruent faces compared to low trust incongruent faces. Given that the two groups of faces are much more polarised in terms of the initial difference in trust ratings we would expect that any interactions with validity should be much stronger than that observed in the previous experiments finding interactions with face gender.

There is also some evidence that trustworthy faces can elicit a stronger gaze-cueing effect than untrustworthy faces (Petrican et al., 2012; Süßenbach & Schönbrodt, 2014). Given that learning of trust judgements is influenced by gaze-cueing, it would be interesting to see whether we can replicate the effects of face trustworthiness on gaze-cueing and whether this would lead trustworthy faces to elicit a larger trust effect in ratings than untrustworthy faces. Indeed, one previous study, found the gaze-cueing effect correlated with the trust effect (Bayliss & Tipper, 2006).

## Methods

### Participants

Participants were 30 students from Bangor University with a mean age of 20 (*SD* = 3), the overall majority of which were female (*N* = 29) and right handed (*N* = 25). All participants were neurologically normal with normal or corrected-to-normal vision and received course credit for taking part. The Bangor University ethics committee approved all procedures.

**Stimuli and Apparatus**

The experiment was displayed on a 19” Iiyama Vision-master CRT display. The face stimuli were taken from the Karolinska database of emotional faces (KDEF; Lundqvist, Flykt & Ohman, 1998) (see Figures 25 & 26). All faces were female and selected based on ratings from Oosterhof and Todorov (2008). Sixteen faces were selected, eight of which were the most trustworthy judged faces (see Figure 25) and eight of which were the most untrustworthy judged faces (see Figure 26). The faces used were AF01, 06, 07, 10, 12, 13, 14, 16, 19, 21, 23, 25, 29, 30, 31, & 33NES. Left and rightward gazing faces were manipulated by moving the iris into the left/ right hand side of the eye. The faces were 300 × 406 pixels in size and the objects were 175 × 175 pixels in size. Stimuli were displayed at a screen resolution of 800 × 600 pixels in the cueing phase and at 640 × 480 pixels in the rating phases. Faces were counterbalanced across participants for congruency.

*Figure 25.* The faces judged to be extremely trustworthy that were used in experiment 4. Taken from the Karolinska database of emotional faces (KDEF Lundqvist et al., 1998) based on ratings from Oosterhof and Todorov (2008).

*Figure 26.* The faces judged to be extremely untrustworthy that were used in experiment 4. Taken from the Karolinska database of emotional faces (KDEF Lundqvist et al., 1998) based on ratings from Oosterhof and Todorov (2008).

**Design and Procedure**

The experimental design was the same as experiment 2. However, instead of half of the faces being male/ female, the faces were either trustworthy or untrustworthy. As no EEG or EMG was recorded, participants were familiarised with the objects, completed the initial rating phase followed by the cueing phase and the final rating phase.

**Data Screening**

Errors (*M* = 3.41%, *SD* = 2.7%) and outliers two standard deviations above or below each participants mean or above 1500 ms (*M* = 7.19%, *SD* = 5.5%) were removed from all analyses, as in previous experiments. Paired-samples t-tests were used to test for differences in the percentage of errors and outliers between congruent and incongruent conditions. There was no significant difference between conditions in the number of errors, $t$ (29) = 1.28, $p$ =

.211, [-.20 .87] or outliers, *t* (29) = -.323, $p = .75$, [-.77 .56]. See Table 8. As in previous experiments, participants were divided into two groups depending whether they did or did not exhibit a trust effect. Overall, 21 participants showed a trust effect and 9 participants did not. Due to the large difference in the numbers of participants in each group and the fact that between-subject variability was not of primary interest in this experiment, we did not enter trust effect as a between-subjects factor in the cueing ANOVA.

Table 1. *Means and standard deviations of the percentage of errors and outliers across conditions.*

| | Errors | Outliers |
|---|---|---|
| Congruent | 1.5(1.3) | 3.5(2.9) |
| Incongruent | 1.9(1.7) | 3.7(2.9) |

## Results

### Trust Ratings

Trust ratings were initially analysed with a 2 × 2 × 2 within-subjects ANOVA with factors comprising face trustworthiness, validity and time of rating. There was a significant effect of face trustworthiness, *F* (1,29) = 123.6, $p < .0001$, $\eta_p^2 = .810$, with high trust faces (*M* = 20.12, *SEM* = 2.93) being judged more trustworthy than low trust faces (*M* = -23.34, *SEM* = 3.45). There was also a significant interaction between validity and time of rating, *F* (1,29) = 11.11, $p = .002$, $\eta_p^2 = .277$, with incongruent faces being rated less trustworthy compared to congruent faces at the end compared to the beginning of the experiment (see Figure 27). However, there was no face trustworthiness × time of rating interaction, *F* (1,29) = .504, $p = .484$, $\eta_p^2 = .017$. Thus, although low trust faces became rated more trustworthy than high trust faces when congruent and high trust faces became judged less trustworthy than low trust faces when incongruent, the interaction did not reach significance as would be evidenced by a face trustworthiness × time of rating interaction. No other effects were significant.

To confirm whether the significant main effects and interactions described above were driven by effects in the beginning ratings, end ratings or both, separate validity × face trustworthiness ANOVAs were run on the beginning and end ratings separately. These analyses showed that at the beginning rating, there was a significant main effect of face trustworthiness, *F* (1,29) = 122.0, $p < .0001$, $\eta_p^2 = .808$, but no effect of validity, *F* (1,29) = 1.29, $p = .27$, $\eta_p^2 = .04$, or a validity × face trustworthiness interaction, *F* (1,29) = .011, $p =$

.919, $\eta_p^2 = .000$. At the end ratings, there was a significant main effect of face trustworthiness, $F(1,29) = 44.8$, $p < .0001$, $\eta_p^2 = .607$, a significant effect of validity, $F(1,29) = 8.08$, $p = .008$, $\eta_p^2 = .22$, but no significant validity × face trustworthiness interaction, $F(1,29) = .001$, $p = .974$, $\eta_p^2 = .000$.

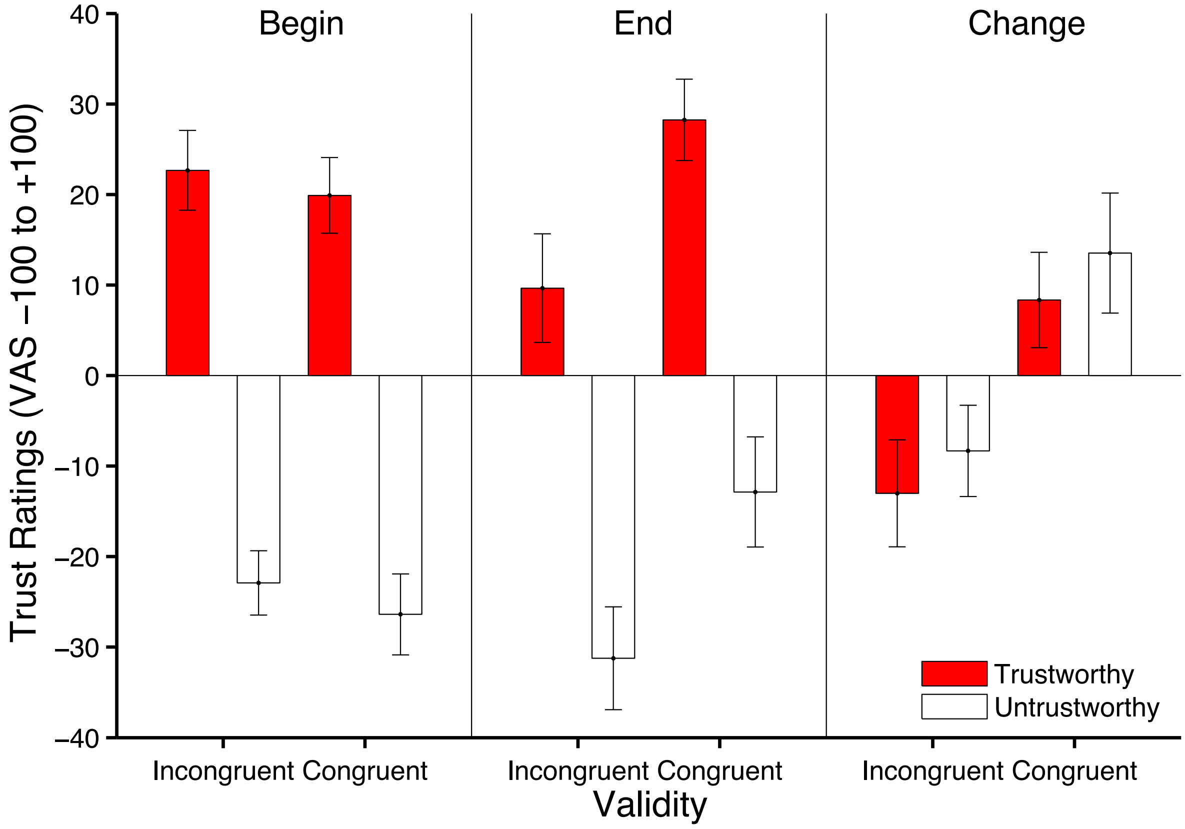


*Figure 27.* Trustworthiness ratings of faces before (left panel) and after (middle panel) cueing and the change in ratings from before to after curing (right panel) in experiment 4. Error bars show +/- 1 standard error.

## Gaze-Cueing Reaction Times

Reaction times were analysed with a 2 × 2 × 5 within-subjects ANOVA with factors of face trustworthiness, validity and block. There was a significant main effect of validity, $F(1,29) = 27.34$, $p < .0001$, $\eta_p^2 = .485$, with reaction times being slower for incongruent faces (*M* = 864.77, *SEM* = 22.7) compared to congruent faces (*M* = 829.38, *SEM* = 22.6) (see Figure 28). There was also a main effect of block, $F(4, 116) = 24.5$ $p < .0001$, $\eta_p^2 = .458$, which was due to a significant linear decrease in reaction times as the experiment progressed, $F(1,29) = 43.6$, $p < .0001$, $\eta_p^2 = .601$. However, there was no block × validity interaction, $F(4, 116) = 1.16$, $p = .334$, $\eta_p^2 = .038$, showing that the effects of validity remained constant throughout the experiment. There was no main effect of facial trustworthiness, $F\ (1, 29) = 1.9$, $p = .181$, $\eta_p^2 = .061$, no face trustworthiness × validity interaction, $F\ (1, 29) = .02$, $p = .897$, $\eta_p^2 = .001$, and no face trustworthiness × validity × block interaction, $F\ (4, 116) = 1.51$, $p = .205$, $\eta_p^2 = .049$. No other main effects or interactions approached significance.

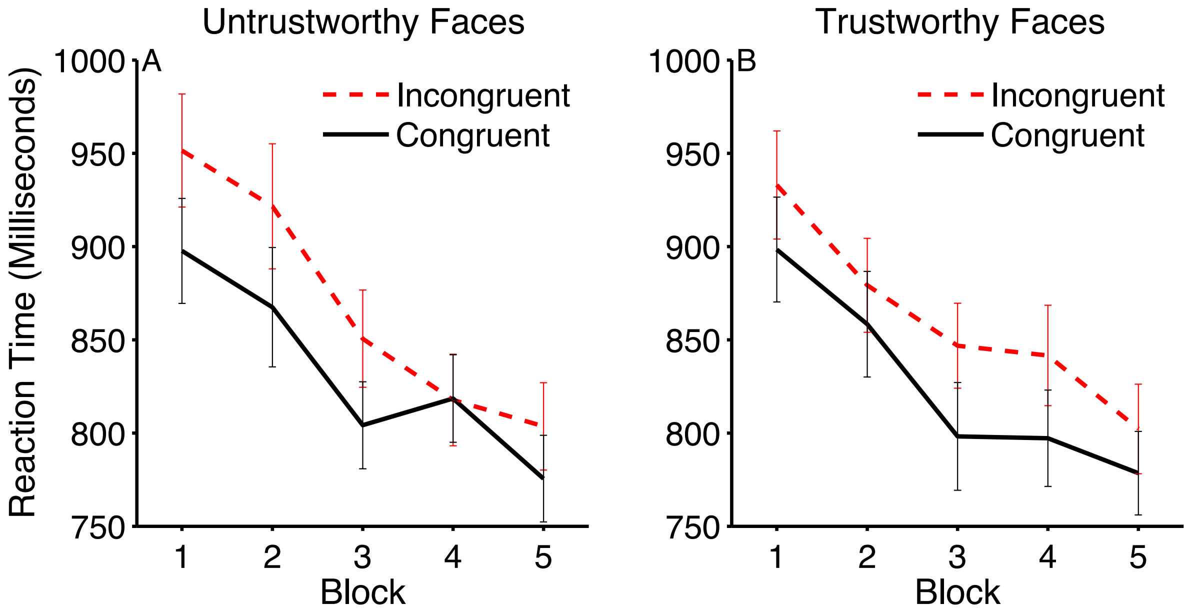


*Figure 28.* Gaze-cueing reaction times for untrustworthy faces (left panel) and trustworthy faces (right panel) in experiment 4. Error bars show +/- 1 standard error.

## Discussion

Experiment 4 investigated the effects of initial trustworthiness based on facial appearance on the learning of trust from gaze-cues. Our hypotheses were based on the findings of experiments 1, 2 and 3 which showed that the learning of trust from congruent gaze-cues was larger for male faces and the learning of mistrust from incongruent gaze-cues was larger for female faces. Such findings suggested that the initial trustworthiness of the face interacts with the learning of trust from gaze-cues. The effects of gaze-cues on trust appeared to be larger when the trustworthiness of the gaze-cue is mismatched with the trustworthiness of the behaviour anticipated based on visual appearance of the face. This could be because trust learning from gaze conforms to prediction-error learning rules (Rescorla & Wagner, 1972). In a stronger test of this hypothesis, we compared learning of trust from faces rated extremely trustworthy or untrustworthy based on appearance.

Whilst replicating the effects of facial appearance on judgements of trustworthiness and effects of validity on the learning of trust, there was no interaction between validity and facial trustworthiness, although the means were in the anticipated direction. The lack of a significant interaction is consistent with previous research investigating the learning of trustworthiness from positive/ negative social cues when emitted by extremely trustworthy or untrustworthy faces. Aguado et al. (2011) paired trustworthy and untrustworthy faces with either happy or angry facial expressions. Although there were main effects of expression and trustworthiness, with trustworthy and happy faces being rated more trustworthy than untrustworthy and angry faces, there was no interaction between expression and trust. Similarly, Rudoy and Paller (2012) found main effects of face trustworthiness in ERP's and trust ratings when faces were paired with positive and negative trait words but no interaction between trait learning and face trust. Chang et al. (2010) found the opposite interaction between face trust and identity-contingent high or low probability of monetary returns in the ultimatum game. After receiving monetary gains or losses from particular faces, trustworthy and high probability faces were rated more trustworthy than untrustworthy and low probability faces. There was also an interaction, as the difference between trustworthy and untrustworthy faces was higher for faces with higher probability of reciprocating.

The findings of the current experiment suggest that the interaction of initial trustworthiness judgements based on facial appearance with the learning of trust from gaze-cues may only occur when the cues to trustworthiness are subtle, as in the case of face gender. This could be similar to the explanation for the lack of effects of exclusion and inclusion observed in experiment 3. If participants are more clearly aware of the source of their emotional reactions to the faces, in this case due to the appearance of the face, they may be less likely to attribute that emotion to the gaze-cues. Indeed, the ratings at the beginning of the experiment showed that trustworthiness was highly polarised between groups. Delgado, Frank and Phelps (2005) investigated the effects of prior knowledge of trustworthiness on the learning of trust. Faces that were previously described as trustworthy, untrustworthy or neutral acted as trustees in a trust game in which all faces had a 50/50 probability of returning a fair or unfair return. Participants showed evidence of learning trustworthiness from trustee outcomes. Before the trust game, trustworthy trustees were rated more trustworthy than neutral trustees and bad trustees were rated less trustworthy than neutral trustees. However, after the trust game, all faces were judged to be of equal trustworthiness. Participants also showed biases by the initial trustworthiness descriptions during the trust game as they chose to invest more often with trustworthy compared to neutral faces and less often with untrustworthy compared to neutral faces. However, fMRI showed the prediction-error reinforcement-learning signal in the striatum to positive compared to negative feedback from the trustees was reduced for both the trustworthy and untrustworthy faces compared to neutral faces. This suggests that when participants can already predict the trustworthiness of the face, based on behavioural descriptions or facial appearance, there is reduced implicit prediction-error learning despite observable changes in trust ratings. This may partially account for the lack of an interaction effect observed in the current experiment.

The lack of updating of trust based on implicit prediction-error learning could also be related to our paradigm. The decisions participants have to make about trust based on facial cues alone at the initial ratings may contribute to the impressions participants form. It has been consistently shown that decisions/ choices can influence preference (Ariely & Norton, 2008). For example, Sharot, Velasquez and Dolan (2010) found that when participants believed they had made blind choices about which of two holiday destinations they preferred, they subsequently preferred those they believed they had chosen, when in reality the destinations chosen were randomly selected by an algorithm. Similarly, participant's ratings may be less amenable to change when participants believe they have made a polarized decision about trust to begin with and when there are salient reasons (e.g. face appearance) for those judgements. It remains to be demonstrated whether the predicted interaction effect would be obtained if the initial trust rating was omitted.

Overall, the findings of experiment 4 could suggest that the learning of trust from gaze-cues is mediated by evaluative conditioning learning rules as opposed to prediction-error learning. The most relevant learning theories would appear to be the referential (Baeyens et al., 1992), holistic (Martin & Levey, 1978; 1994) and implicit misattribution accounts (Jones et al., 2009). As such, participants may not be learning a predictive association between face identity and gaze-cues. Instead, participants may be forming referential connections between the two sources of information, according to Hebbian (1949) learning, or integrating both into the same representation. The affective properties of the gaze-cues may then diffuse to face identity producing changing in ratings. Implicit misattribution may also be at play and may explain the interactions between face gender and validity observed in experiments 1, 2 and 3. If participants implicitly base their trust judgements on facial appearance and this is not a salient source of their evaluation it may become mixed with information from other sources, such as from gaze-cues.

Surprisingly we did not observe any interactions between trustworthiness and validity in the gaze-cueing reaction times in spite of several studies to the contrary. For example, high social status (Dalmaso, Pavan, Castelli & Galfano, 2011) and trustworthy faces (Petrican et al., 2012; Süßenbach & Schönbrodt, 2014) have been shown to elicit a larger gaze-cueing effect than low social status and untrustworthy faces. However, the effects for trustworthiness of faces were only observed in older participants with high cognitive control and in low anxious participants, respectively. Similarly, Hudson, Nijboer and Jellema (2012) found that participants scoring low, as opposed to high, on autism traits, showed a greater gaze-cueing effect for pro-social avatars who had previously smiled and gazed toward participants or frowned looking away from participants compared to anti-social avatars who looked away from participants when smiling and frowned when looking towards. We did not divide our participants on these variables, which may partly explain why no effects were observed.

In conclusion, this experiment has investigated the interaction between face trustworthiness and learning of trust from gaze-cues. We predicted that the learning of trust from congruent cues would be larger for untrustworthy faces and the learning of mistrust from incongruent cues would be larger for trustworthy faces. This is based on the assumption that emotional reactions and learning are greater when there is a mismatch between how trustworthy participants expect the face to behave and how trustworthy the face actually behaves. Although the predicted interaction effect was in the anticipated direction, it did not reach significance. This could be because participants can more readily attribute their emotional responses to facial appearance and gaze-cues when they are more clearly discriminable, leading to decreased interaction between the two. It could also be because the decisions made about trust based on facial appearance at the initial rating reduce implicit prediction-error learning. Therefore, a stronger test of the hypothesis that the initial trustworthiness of the face interacts with the learning of trust may be to omit the initial trust rating phase. It may also be more clearly elucidated at the neural level.

# Chapter 6

## Experiment 5: The specificity of the learning of evaluations of faces from gaze-cues and the mediating role of emotion. A comparison with non-social arrow-cues

All of the previous experiments conducted up to now have investigated the learning of trustworthiness from gaze-cues and the mediating role emotion may play in this process. However, it is unclear whether the learning of evaluations is specific to the social qualities of faces, gaze and trust or whether it is reflective of a domain general relationship between the validity of attentional cues, emotion and evaluation. If the latter were the case then identity-contingent arrow cues would also be expected to generate a similar change in affective states and evaluations as such stimuli cue attention in a very similar manner to gaze (Bayliss & Tipper 2005; Ristic, Friesen & Kingstone, 2002; Tipples, 2002; 2008).

However, unlike arrows, faces and gaze are highly social and are characterised by several distinct channels of information that are motivationally relevant and emotionally evocative. This includes facial expression, gaze direction and identity, which have specific neural substrates. Unlike non-social arrows, other people also have mental states, such as intentions, which may be used as information when interpreting and responding emotionally to others' behaviour. Indeed, mental state information has even been shown to modulate gaze-cueing effects (Teufel et al., 2010). Faces are also more interesting to participants as they are fixated first and more often than arrows presented concurrently in natural scenes (Birmingham, Bischof & Kingstone, 2009). This may be why infants tend to overtly follow the gaze of faces that consistently gaze congruently more than faces that consistently gaze incongruently, an effect that is absent for identity-contingent congruent/ incongruent arrow cues (Tummeltshammer et al., 2014). Therefore it may be predicted that identity-contingent gaze-cues are a stronger elicitor of emotion and lead to greater learning of evaluations than arrow cues. Consistent with this reasoning is research showing that liking of objects is increased when faces consistently gaze toward them in the cueing paradigm compared to when consistently gazed away from (Bayliss et al., 2013; Capozzi, Bayliss, Elena & Becchio, 2014; Manera, Elena, Bayliss & Becchio, 2014; Treinen, Corneille & Luypaert, 2012), but this effect of validity on object liking is absent when the cues are arrows or hands (Bayliss, Paul, Cannon & Tipper, 2006; Ulloa, Marchetti, Taffou & George, 2015).

In experiment 5, we compare the learning of evaluations from gaze and arrow cues and their relation to emotion-related EMG activity in order to determine whether the effects observed in previous experiments are specific to the social qualities of faces and gaze or are a product of a domain general mechanism that is dependent only upon attentional cueing. To this end, arrow stimuli were created that had distinctive fractal patterns and colours in order to be differentiated from one another and have a memorable identity to a similar degree as the faces. To double check that this is the case we performed a preliminary visual search and memory experiment in which participants had to detect face or arrow target stimuli in the presence of distractors (Müller & Krummenacher, 2006). Afterwards, participants' memory was tested using a task in which they had to recognise which faces or arrows had been presented in the previous visual search task amongst faces and arrows that were not present.

In separate gaze cueing and arrow cueing experiments we measured liking ratings instead of trust. This is because evaluations of trustworthiness do not readily generalize to arrow stimuli, which clearly do not have personality traits. In contrast, liking can apply to both types of stimuli and is related to trustworthiness ratings of faces, although the latter dimension is more strongly related to face evaluation (Oosterhof & Todorov, 2008). The investigation of whether identity-contingent gaze-cues can influence liking judgements in a similar manner as it does trust is also interesting in its own right. The lack of an effect or a weaker effect compared to previous studies would suggest a special relationship between faces, gaze and trust. There is some evidence for this. Although one experiment has shown effects of identity-contingent gaze-cues on preference judgements (Bayliss et al., 2013), this effect was not replicated in others (Bayliss et al., 2013; Bayliss & Tipper, 2006) whereas the learning of trust from gaze-cues is stronger and more consistent.

## Preliminary Visual Search and Memory Experiment

### Methods

#### Participants

There were 25 participants in total of which 16 were female and 22 right handed. The mean age of participants was 19 (*SD*=1). Participants were neurologically normal with normal or corrected-to-normal vision. All procedures were reviewed and granted by the ethics review board at Bangor University.

#### Stimuli and Apparatus

There were 16 arrow stimuli and 16 face stimuli. The face stimuli were the same as in experiments 2 and 3. The arrow stimuli consisted of a shaft, an arrow head and a tail. As during gaze-cueing a direct face begins and ends the trial, the same logic was applied to the arrows. Thus, during cueing, each trial began and ended with the head and tail straight in an H shape. Only at the same time point of the gaze cue did the shaft shift to the left or right and the head and tails bend into their respective shapes (see Figure 32). The visual search task was performed on the arrows and faces in their H/ direct gaze state. Half of the arrows had a red border and the other half had a blue border. This was to create two arbitrary categories that emulate the gender category of faces. Within the borders of the arrow was a distinctive fractal pattern (see Figure 29). For the visual search task, the arrows were sized at 385 × 295 pixels and the faces were 300 × 385 pixels. The experiment was run using E-Prime 1.0 (Psychology Software Tools, Inc.) on a 24” Samsung SyncMaster BX2431 LED display with a screen resolution of 1920 × 1080, which was 569 × 342 mm in dimensions and had a 500 Hz refresh rate. The faces and arrows were divided into two groups half of which were male/ female or blue/ red. For faces this was the same as the groups used in the previous experiments.

*Figure 29.* The 16 arrow stimuli used in experiment 5 in their direct state.

### Design and Procedure

Participants initiated each trial by pressing space bar after which a fixation cross appeared for 1500 ms. Then a face or arrow appeared in the centre of the screen and remained for 2000 ms. This was the target. When the target disappeared, participants saw another fixation cross for 1500 ms and then an array of four faces or arrows in each quadrant of the screen. On half of the trials the target was present within the array whereas on the other half the target was not present (see Figure 30). Participants were instructed to press the "U" button if the target was present and the "N" button if the target was not present as quickly but as accurately as possible. The response keys were counterbalanced across participants. Afterwards, a blank screen was presented for 1000 ms. If participants responded incorrectly or did not respond within 3500 ms, the trial was treated as an error and an error tone was played for 1000 ms. Two versions of the task were created each containing half of the faces /arrows. Group A stimuli were in one experiment and group B in a counterbalanced version. The group not used in the visual search task were used as distractors in the subsequent memory test. There were 4 blocks in total each containing 48 trials. Each of the faces/ arrows was presented as the target on 6 trials each, on half of which the target was present and on the other half not present. The order of trials, the identity and position of the distractors were all randomised within blocks. Before the experiment, participants completed 8 practice trials.

In the subsequent memory test immediately after the visual search task, all 16 arrows/ faces were presented to the participant and they were required to state whether they remembered seeing the arrow/ face in the previous task. On each trial participants pressed space to begin, saw a fixation cross for 1000 ms followed by a stimulus in the centre of the screen. Participants were asked *"Did you see this shape (face) in the previous task?"* At the bottom left and right hand side of the screen were boxes labelled "*No*" and "*Yes*", respectively. Participants clicked the box to indicate whether they remembered seeing the stimulus. They were then asked to give a rating of how confident they were that this was correct, rating the stimuli from "*Not sure at all*" to "*Absolutely certain"* (see Figure 30). Each participant undertook both the visual search/ memory task with arrows and faces. The order in which each experiment was completed was counterbalanced across participants.

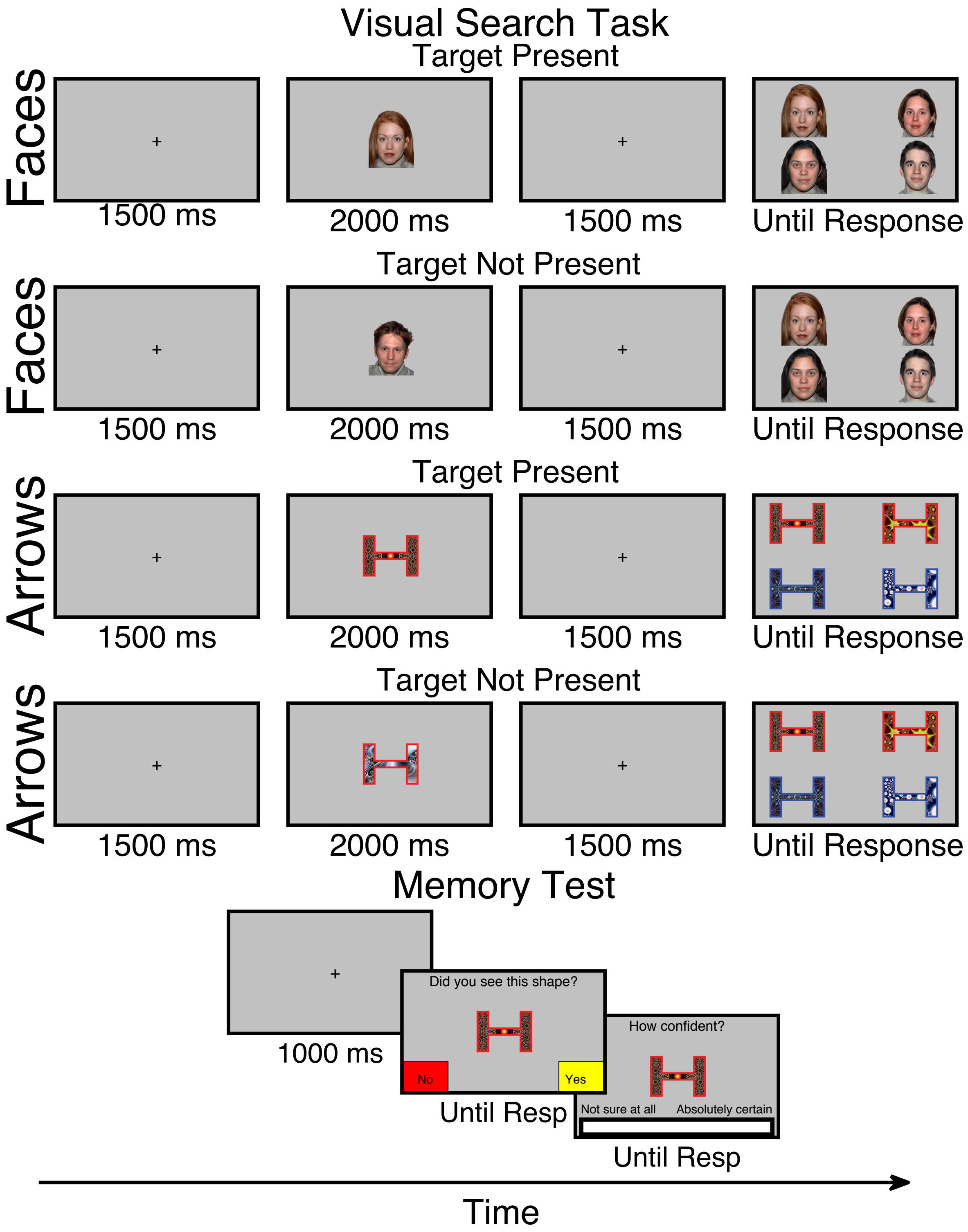


*Figure 30.* Trial procedure for the visual search and memory task with faces and arrows. In the visual search task participants pressed space bar to initiate each trial. A fixation cross appeared in the centre of the screen for 1500 ms followed by a target stimulus for 2000 ms. After another fixation for 1500 ms, participants were presented with the search array. On half of the trials the target was present within the array and on the other half the target was not present. The trial terminated when participants made a key press to indicate whether the target was present or not present. For the memory test, participants were presented with each

stimulus in the centre of the screen and clicked “yes” or “no” to indicate whether they remembered seeing the stimulus in the visual search task. This was followed by a confidence rating. Half of the 16 faces/ arrows presented in the memory test had been presented in the visual search task whereas the other half had not been presented and were therefore distractors. Exactly the same procedure was used with the face stimuli.

### Data Screening

All trials on which participants made an error or did not respond in the allotted time were removed along with reaction time outliers that were two standard deviations above or below each participants mean. A 2 × 2 within-subjects ANOVA was run on the errors and outliers with factors of stimulus type and target presence. In the outliers, there was a significant main effect of presence, $F\,(1,24) = 16.69, p < .0001, \eta_p^2 = .410$, in that there were more outliers when the target was present ($M = 2.5\%$, $SEM = .16\%$) compared to when the target was not present ($M = 1.58\%$, $SEM = .123\%$). There was also a main effect of stimulus type, $F\,(1,24) = 5.094, p = .033, \eta_p^2 = .175$, as there were more outliers for faces ($M = 2.45\%$, $SEM = .208\%$) compared to arrows ($M = 1.68\%$, $SEM = .156\%$). However, there was no significant interaction between stimulus type and presence, $F\,(1,24) =1.82, p = .19, \eta_p^2 = .071$. These findings are consistent with the reaction time findings described below. In the errors, there was a significant main effect of presence, $F\,(1,24) = 25.17, p < .0001, \eta_p^2 = .512$, with errors being more frequent on present trials ($M = 3.75\%$, $SEM = .561\%$) compared to not present trials ($M = 1.68\%$, $SEM = .32\%$). However, there was no significant effect of stimulus type, $F\,(1,24) = .657, p = .426, \eta_p^2 = .027$, and no significant interaction between stimulus type and presence, $F\,(1,24) = 2.72, p = .112, \eta_p^2 = .102$. The overall difference in the percentage of trials removed from each condition was low (see Table 9).

Table 9. *Means and standard deviations of the percentage of errors and outliers across conditions of the visual search task.*

| | Faces | | Arrows | |
|---|---|---|---|---|
| | Present | Not Present | Present | Not Present |
| Errors | 3.9(3.20) | 1.35(1.44) | 3.6(2.76) | 2.0(2.03) |
| Outliers | 1.8(1.16) | 2.98(1.47) | 1.35(.62) | 2.7(1.29) |

## Results

The reaction times from the visual search task were analysed with a 2 × 2 × 4 repeated measures ANOVA with factors of stimulus type, target presence and block. There was a significant main effect of presence, *F* (1,24) = 47.24, $p < .0001$, $\eta_p^2 = .663$, showing that participants were faster on target present trials (*M* = 853, *SEM* = 25.5) compared to target non-present trials (*M* = 941.9, *SEM* = 33.3), the typical visual search effect (see Figure 31). There was also a significant effect of stimulus type, *F* (1,24) = 6.21, $p = .02$, $\eta_p^2 = .206$, as arrows (*M* = 872.51, *SEM* = 32.8) were distinguished more quickly than faces (*M* = 923, *SEM* = 28.4). There was no main effect of block, *F* (3, 72) = 2.4, $p = .075$, $\eta_p^2 = .091$, or any interactions with stimulus type, *F* (3, 72) = 1.98, $p = .137$, $\eta_p^2 = .073$ and presence, *F* (3, 72) = .475, $p = .7$, $\eta_p^2 = .019$. However, more importantly, there was no significant target presence × stimulus type interaction, *F* (1,24) = 1.175, $p = .289$, $\eta_p^2 = .047$, showing that visual search performance was comparable across the two types of stimuli. There was also no presence × stimulus type × block interaction, *F* (3, 72) = .559, $p = .644$, $\eta_p^2 = .023$. For the memory test, hits and correct rejections were scored as 1 and misses and false alarms were scored as -1 and summated across conditions/ participants. All participants performed at ceiling (16/16). This shows that both types of stimuli are equivalent in perceptual distinctiveness and memorability. Therefore, any differences between arrows and faces in the learning of liking and EMG activity in response to cueing can be confidently rejected as due to differences in terms of how the different stimuli are processed visually or remembered.

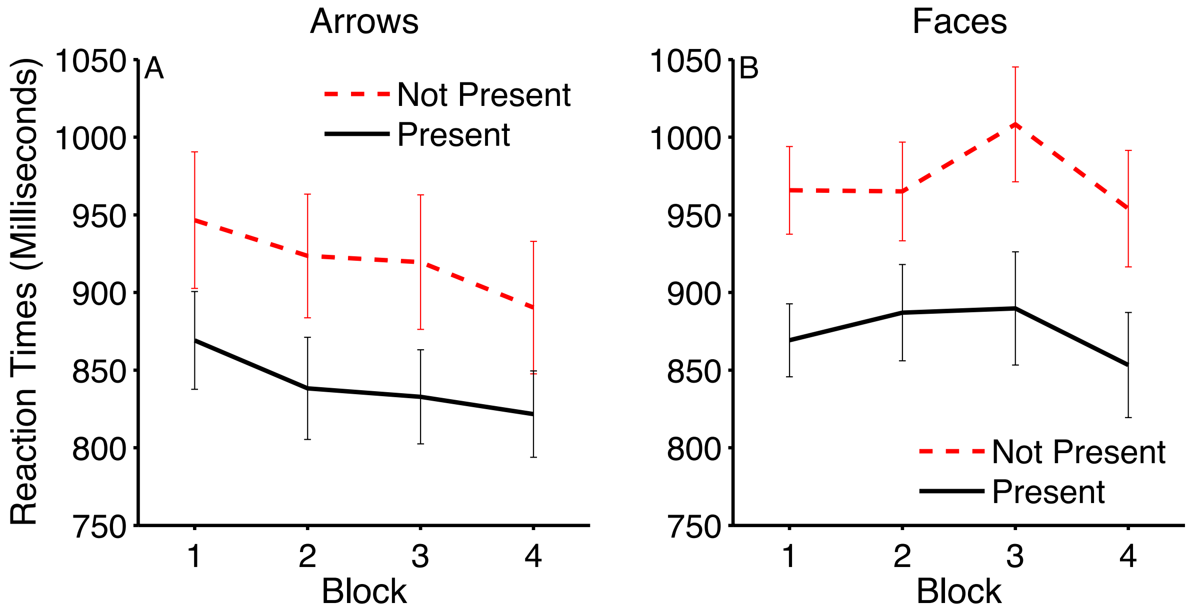


*Figure 31.* Visual search reaction times taken to detect whether arrow (let panels) and face (right panels) targets were present (solid line) or not present (dashed line). Shown across blocks. Error bars show +/- 1 standard error of the mean.

## Arrow- and Gaze-Cueing Experiments

### Methods

#### Participants

In the arrows group there were 31 participants in total of which 29 were right handed and 21 were female. Participants had a mean age of 20 (*SD* = 3.7). In the faces group there were also 31 participants of which 25 were right handed and 26 were female. The mean age of participants was again 20 (*SD* = 3.4).

#### Stimuli

The arrows were 660 × 424 pixels in size whereas the faces were 504 × 600 and the objects 350 × 263. The experiment was run with a screen resolution at 1920 × 1080 on a 24" Samsung SyncMaster BX2431 LED display, which was 569 × 342 mm in dimensions and had a 500 Hz refresh rate.

#### Design and Procedure

Two versions of the experiment were run on separate groups of participants. One of the experiments used arrows as the cueing stimuli (see Figure 32) and a second experiment used faces/ gaze as the cueing stimuli. All procedures were the same as in experiments 2 and 4 except that liking ratings were measured instead of trust ratings. Whereas trust only applies to faces, liking can be used to evaluate both faces and non-faces, such as arrows.

#### Data Screening and Analysis

Differences in errors, outliers and artefacts between conditions in the cueing phase were analysed with a 2 × 2 mixed ANOVA with the within-subjects factors of validity and the between-subjects factor of stimulus type. In terms of errors, there was no significant effect of validity, $F(1, 60) = .22$, $p = .641$, $\eta_p^2 = .044$, stimulus type, $F(1, 60) = .034$, $p = .854$, $\eta_p^2 = .001$, or an interaction, $F(1, 60) = .112$, $p = .739$, $\eta_p^2 = .002$. In terms of outliers, there was no significant effect of validity, $F(1, 60) = 3.28$, $p = .075$, $\eta_p^2 = .052$, stimulus type, $F(1, 60) = 1.27$, $p = .265$, $\eta_p^2 = .021$, or an interaction, $F(1, 60) = .339$, $p = .563$, $\eta_p^2 = .006$. For corrugator artefacts, there was no significant effect of validity, $F(1, 60) = .943$, $p = .335$, $\eta_p^2 = .015$, stimulus type, $F(1, 60) = .429$, $p = .515$, $\eta_p^2 = .007$, or their interaction, $F$

(1, 60) = .72, $p$ = .399, $\eta_p^2$ = .012. For zygomaticus artefacts, there was no significant effect of validity, $F$ (1, 60) = 3.66, $p$ = .06, $\eta_p^2$ = .058, stimulus type, $F$ (1, 60) = .666, $p$ = .418, $\eta_p^2$ = .011, or their interaction, $F$ (1, 60) = .373, $p$ = .544, $\eta_p^2$ = .006. See Table 10. Overall, there were 5 participants who showed a liking effect and 26 participants who did not show a liking effect in the arrows group and 16 participants who showed a liking effect and 15 participants who did not show a liking effect in the faces group.

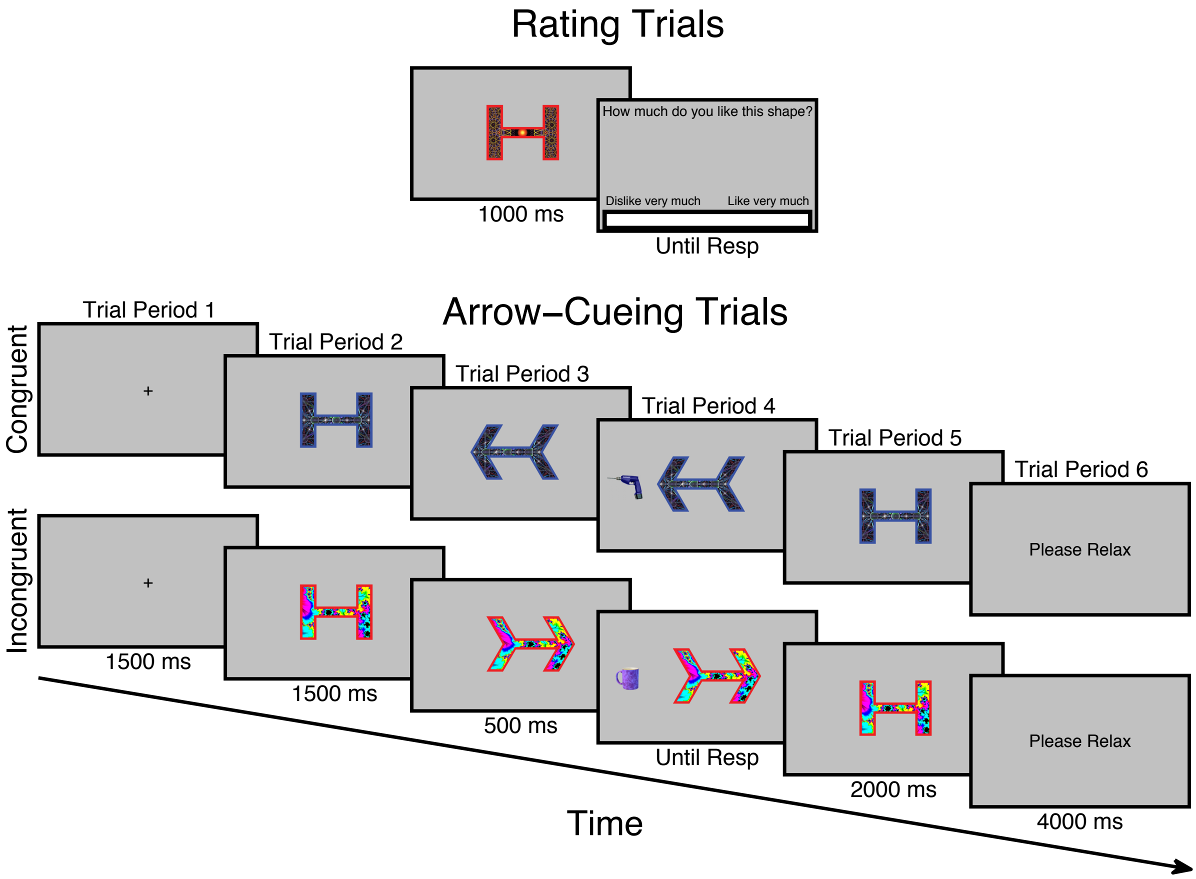


*Figure 32.* Trial procedure on rating and cueing trials for experiment 2 using arrows. On rating trials before and after cueing, participants observed each arrow/ shape for 1000 ms after which a visual analogue rating scale appeared requiring participants to click the point on the scale that represented how much they liked the arrow/ shape. During cueing trials, participants saw a fixation cross for 1500 ms, followed by an arrow/shape in its neutral state for 1500 ms after which it changed shape to point to the left or right for 500 ms when an object appeared to the left or right hand side and disappeared when the participant responded. This also triggered the arrow to go back to its neutral state for another 2000 ms.

Table 10. *Means and standard deviations of percentages of errors, outliers and artefacts across conditions removed from cueing analyses.*

| | Faces | | Arrows | |
|---|---|---|---|---|
| | Incon | Con | Incon | Con |
| Errors | 1.6(1.5) | 1.6(1.7) | 1.7(1.49) | 1.6(1.7) |
| Outliers | 5.4(3.2) | 4.7(2.9) | 6.3(4.9) | 5.9(4.5) |
| Corrugator Artefacts | 2.2(1.64) | 2.2(1.2) | 1.9(1.14) | 2.18(1.48) |
| Zygomaticus Artefacts | 4.8(2.4) | 4.1(2.4) | 4.1(2.07) | 3.8(2.3) |

## Results

### Evaluative Ratings of Liking

The liking ratings were analysed with a 2 × 2 within-subjects ANOVA separately for arrows and faces with factors of time of rating and validity. We did not include face gender or arrow colour as factors in any ANOVAs as these two arbitrary categories are very dissimilar perceptually. For faces, there was a significant interaction between time and validity, $F$ (1, 30) = 7.7, $p$ = .009, $\eta_p^2$ = .204. This appears to be driven by congruent faces as such faces were more liked after cueing ($M$ = 4.2, $SEM$ = 4.2) compared to before cueing ($M$ = -4.1, $SEM$ = 3.6), whereas incongruent faces did not appear to show a change over time (see Figure 33). No other effects approached significance. In contrast, there was no significant validity × time interaction for the arrows, $F$ (1, 30) = .94, $p$ < .34, $\eta_p^2$ = .03, although there was a main effect of time, $F$ (1, 30) = 13.4, $p$ = .001, $\eta_p^2$ = .31, with ratings being more negative after ($M$ = -3.5, $SEM$ = 3.4), compared to before ($M$ = 6.8, $SEM$ = 3.2), cueing.

### Gaze-Cueing Reaction Times

The gaze-cueing reaction times were analysed with a 2 × 2 × 5 mixed ANOVA with the between-subjects factor of stimulus type and the within-subjects factors of validity and block. There was a significant effect of validity, $F$ (1, 60) = 38.39, $p$ < .0001, $\eta_p^2$ = .390, with incongruent trials ($M$ = 871.09, SEM = 21.6) being responded to slower than congruent trials ($M$ = 838.7, $SEM$ = 21.6) (see Figure 34). There was also a significant effect of block, $F$ (1, 60) = 48.95, $p$ < .0001, $\eta_p^2$ = .449, which was significant as a linear trend, $F$ (1, 60) = 94.6, $p$ < .0001, $\eta_p^2$ = .612. However, crucially there was no significant validity × stimulus interaction, *F* (1, 60) = .02, $p = .888$, $\eta_p^2 = .000$, or validity × block × stimulus interaction, *F* (1, 60) = .97, $p = .423$, $\eta_p^2 = .016$. There was also no main effect of stimulus, *F* (1, 60) = 1.3, $p = .26$, $\eta_p^2 = .021$. There was, however, a significant block × validity interaction, *F* (1, 60) = 2.5, $p = .045$, $\eta_p^2 = .04$. This was more significant as a linear trend, *F* (1, 60) = 8.0, $p = .006$, $\eta_p^2 = .118$, and appears to be due to a reduced cueing effect in the final block compared to the other blocks. Although there was no interaction with stimulus type it appears that this is primarily driven by the arrow-cues. For the gaze-cueing experiment, all analyses were reproduced with liking effect as the between-subjects factor. However, no additional main effects or interactions approached significance.

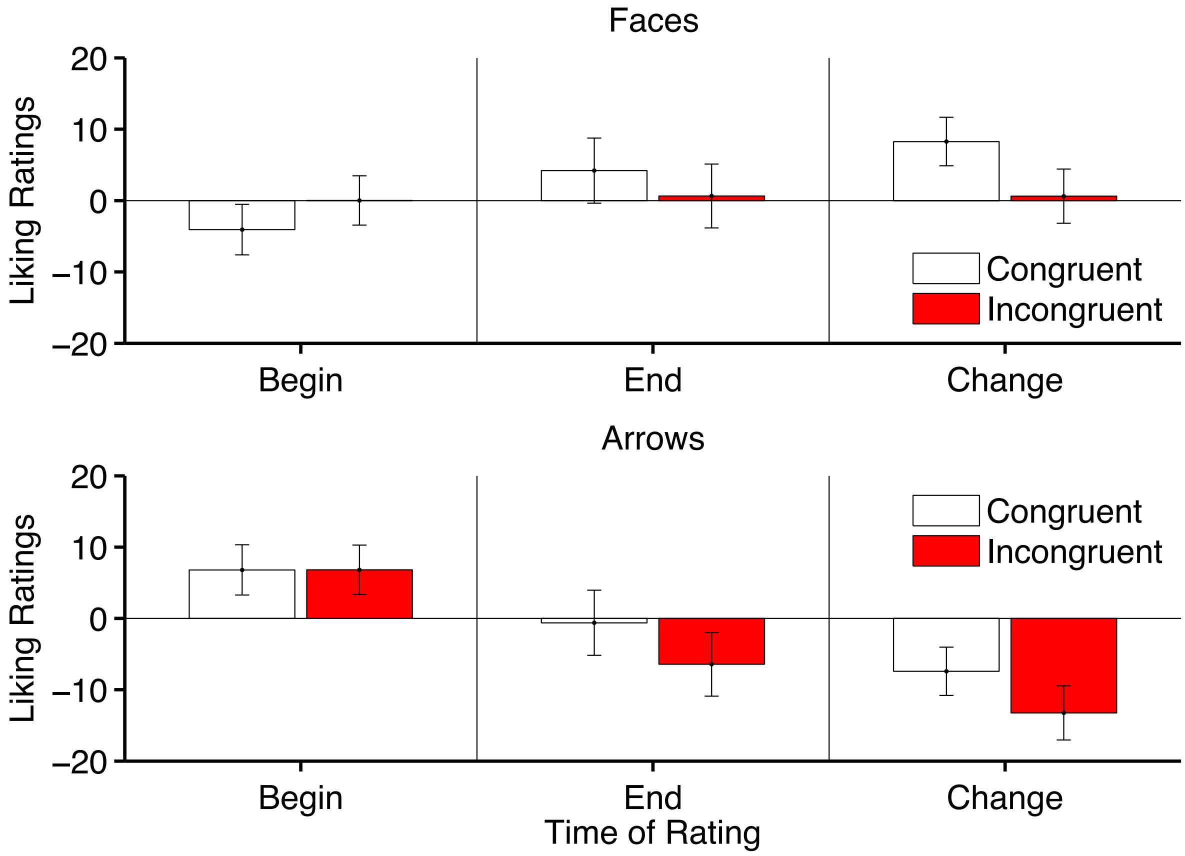


*Figure 33.* Liking ratings of faces (top panel) and arrows (bottom panel) before (left panels) and after cueing (middle panels) and the change in liking from before to after cueing (right panels). Error bars show +/- 1 standard error of the mean.

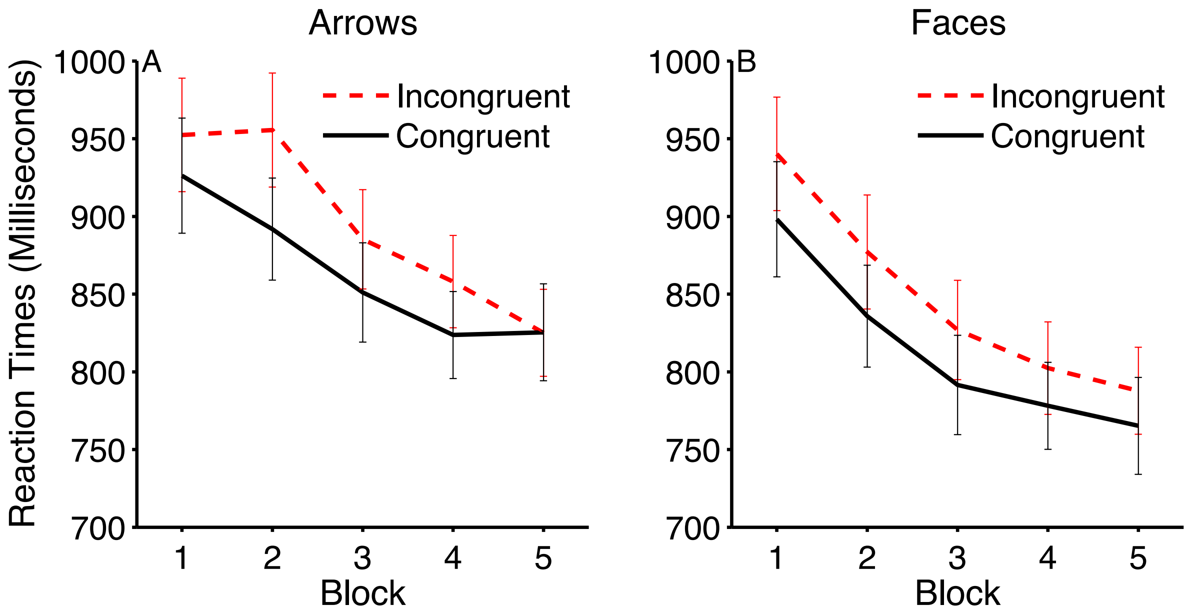


*Figure 34.* Reaction times across blocks for congruent (solid line) and incongruent (dashed line) arrows (left panels) and faces/ gaze (right panels). Error bars show +/- 1 standard error of the mean.

**Facial Electromyography**

The EMG data was analysed in the same manner as in experiments 2 and 3. Trial periods 2 and 3 were analysed using a 2 × 2 × 4 mixed ANOVA with the between-subjects factor of stimulus type and the within-subjects factors of validity and block. Trial periods 4 and 5 were analysed with a 2 × 2 × 5 mixed ANOVA that included all 5 blocks. The factors of face gender and arrow colour were also dropped from EMG analyses for the same reasons as in the previous analyses and also because these variables were not of primary interest. In order to determine whether main effects or interactions were driven by one or both types of stimulus, separate within-subjects validity × block ANOVAs were run in the arrows and faces group separately. All analyses were repeated with liking effect as the between-subjects factor. However, as there were only a small number of participants who showed a liking effect in the arrows group, this analysis was only performed on the faces group data. As there were no additional significant effects of liking, we refrain from reporting these results.

**Corrugator**

In trial period 2 there was a significant main effect of stimulus type, $F\ (1, 60) = 4.23$, $p = .044$, $\eta_p^2 = .066$. This was due to larger activity to arrows compared to faces. In trial period 4, there was a significant main effect of validity, $F\ (1, 60) = 17.43, p < .0001$, $\eta_p^2 =$

.225, due to larger activity in response to incongruent compared to congruent cues. Separate within-subjects ANOVAs in each group showed that the effect of validity in the corrugator was significant in both the arrows group, $F$ (1, 30) = 8.76, $p = .006$, $\eta_p^2 = .225$, and faces group, $F$ (1, 30) = 9.8, $p = .004$, $\eta_p^2 = .25$. In trial period 5, there was a borderline significant main effect of validity, $F$ (1, 60) = 3.98, $p = .051$, $\eta_p^2 = .062$. This was qualified by a significant interaction between stimulus type and validity, $F$ (1, 60) = 8.7, $p = .005$, $\eta_p^2 = .126$. Separate ANOVAs in the faces and arrows group showed that the effect of validity was significant in the faces group, $F$ (1, 30) = 9.75, $p = .004$, $\eta_p^2 = .245$, but not in the arrows group, $F$ (1, 30) = .762, $p = .39$, $\eta_p^2 = .025$. See Figures 35 and 36.

### Zygomaticus

In trial period 2 there was a significant main effect of stimulus type, $F$ (1, 60) = 5.65, $p = .021$, $\eta_p^2 = .086$, due to larger activity to arrows compared to faces. In trial period 3, there was also a significant main effect of stimulus type, $F$ (1,60) = 6.67, $p = .012$, $\eta_p^2 = .10$. Again, this was due to larger activity to arrows compared to faces. There was also a significant interaction between validity, stimulus type and block, $F$ (3, 180) = 3.24, $p = .023$, $\eta_p^2 = .051$. This appears to be driven by larger activity to congruent faces in block 4 and a reversal of the effect for the arrows. Separate ANOVAs in each group showed that the validity × block interaction was not significant in either the arrows group, $F$ (3, 90) = 1.32, $p = .27$, $\eta_p^2 = .042$, nor the faces group, $F$ (3, 90) = 1.7, $p = .17$, $\eta_p^2 = .054$. In trial period 4, there was a significant main effect of validity, $F$ (1, 60) = 15.97, $p < .0001$, $\eta_p^2 = .209$. This was due to larger activity in response to incongruent compared to congruent cues. Separate within-subjects ANOVAs in each group showed that the effect of validity was significant for both the arrows, $F$ (1, 30) = 9.84, $p = .004$, $\eta_p^2 = .247$, and faces group, $F$ (1, 30) = 6.64, $p = .015$, $\eta_p^2 = .181$. In trial period 5 there was a significant validity × stimulus × block interaction, $F$ (4, 240) = 2.6, $p = .043$, $\eta_p^2 = .04$. This effect appears to be due to the same pattern as in trial period 3. Separate ANOVAs in each group showed that the validity × block interaction was not significant in either the arrows group, $F$ (4, 120) = 1.36, $p = .28$, $\eta_p^2 = .041$, nor the faces group, $F$ (4, 120) = 1.98, $p = .10$, $\eta_p^2 = .062$. See Figures 37 and 38.

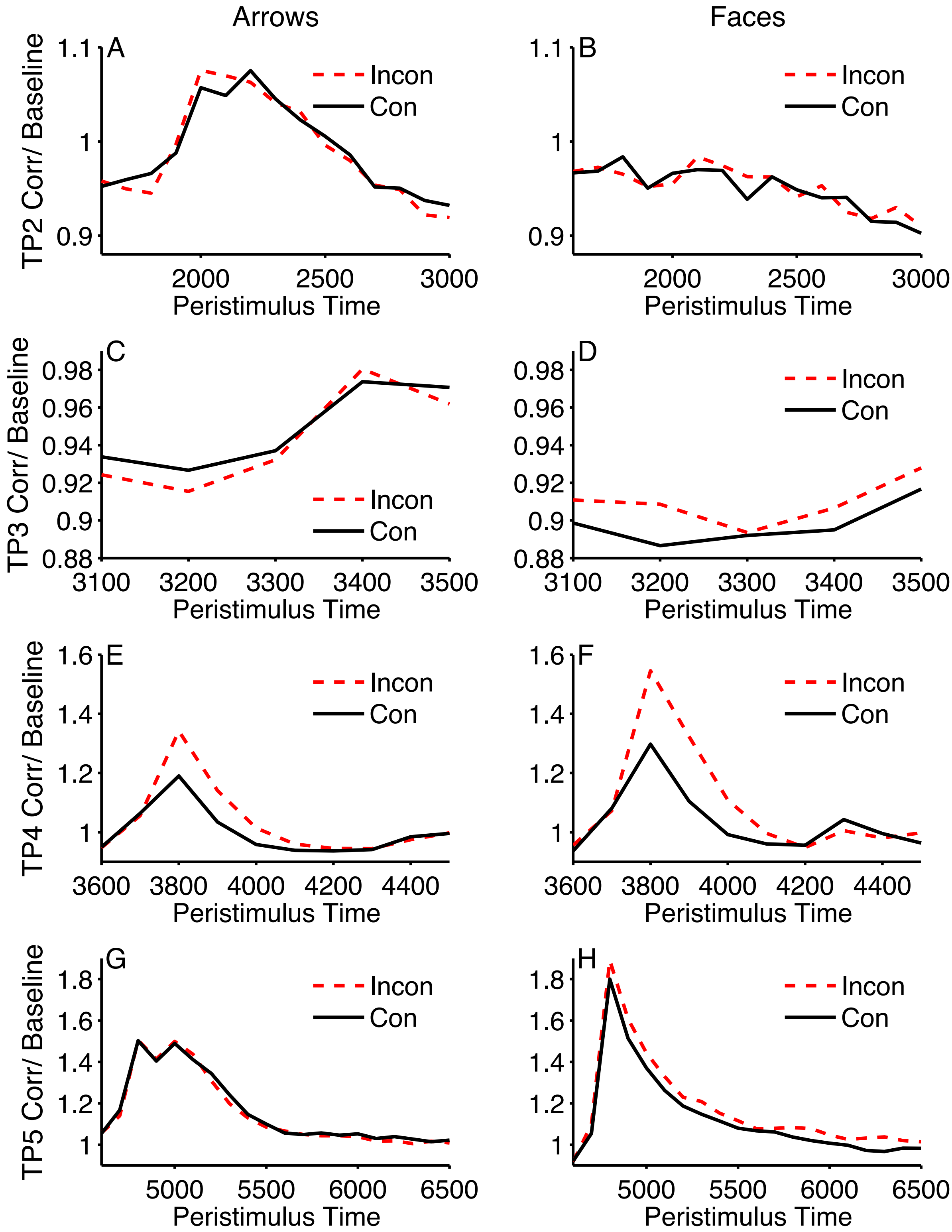


*Figure 35.* Mean stimulus-locked corrugator activity on congruent (solid line) and incongruent trials (dashed line) for arrows (left panels) and faces (right panels) across trial periods 2, 3, 4 and 5 (rows). EMG units on the Y-axis represent the ratio of activity relative to baseline (fixation).

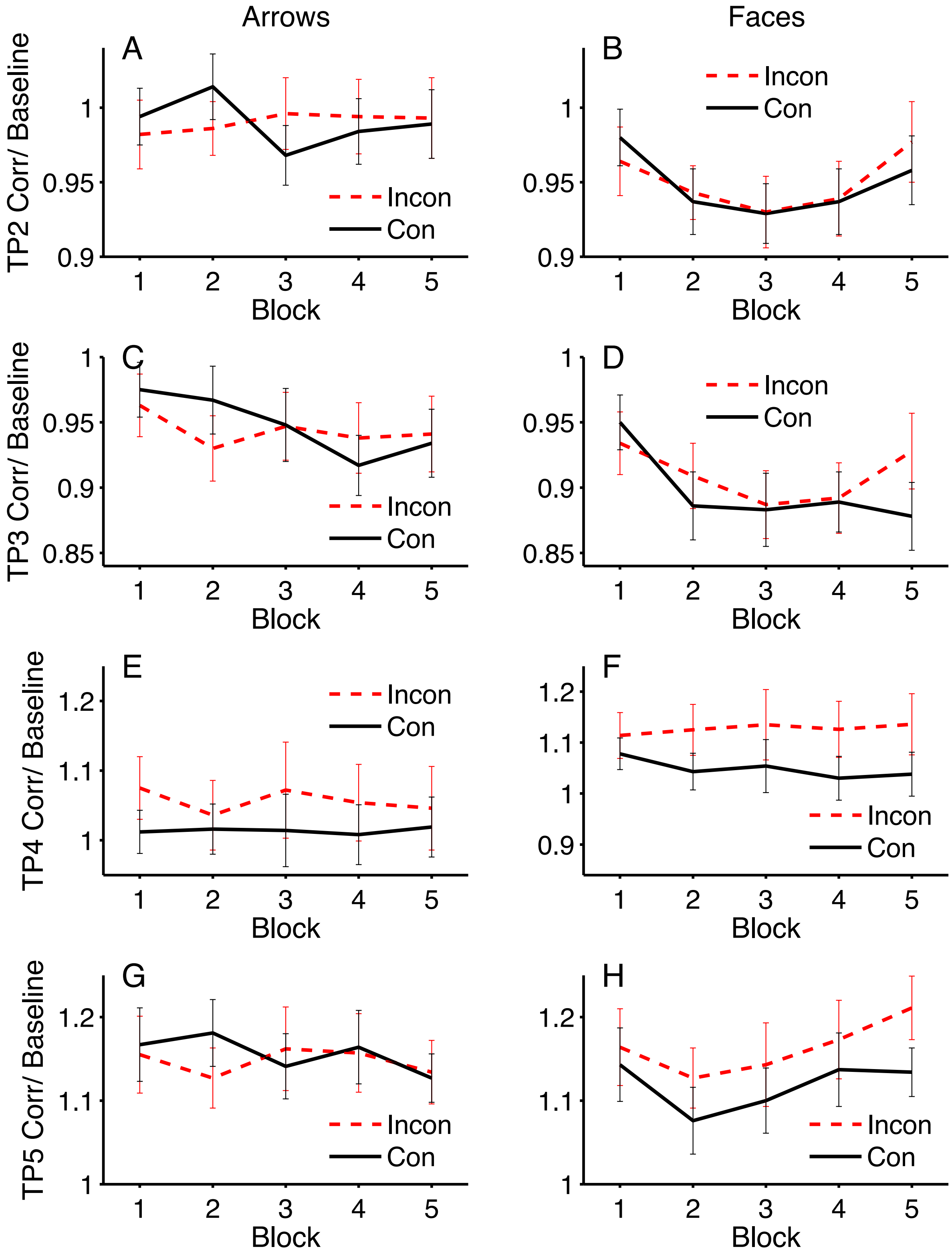


*Figure 36.* Mean corrugator activity across blocks for arrows (left panels) and faces (right panels) in trial periods 2, 3, 4 and 5 (rows). Dashed lines represent incongruent and solid lines represent congruent. EMG units on the Y-axis represent the ratio of activity relative to baseline (fixation). Error bars show +/- 1 standard error of the mean.

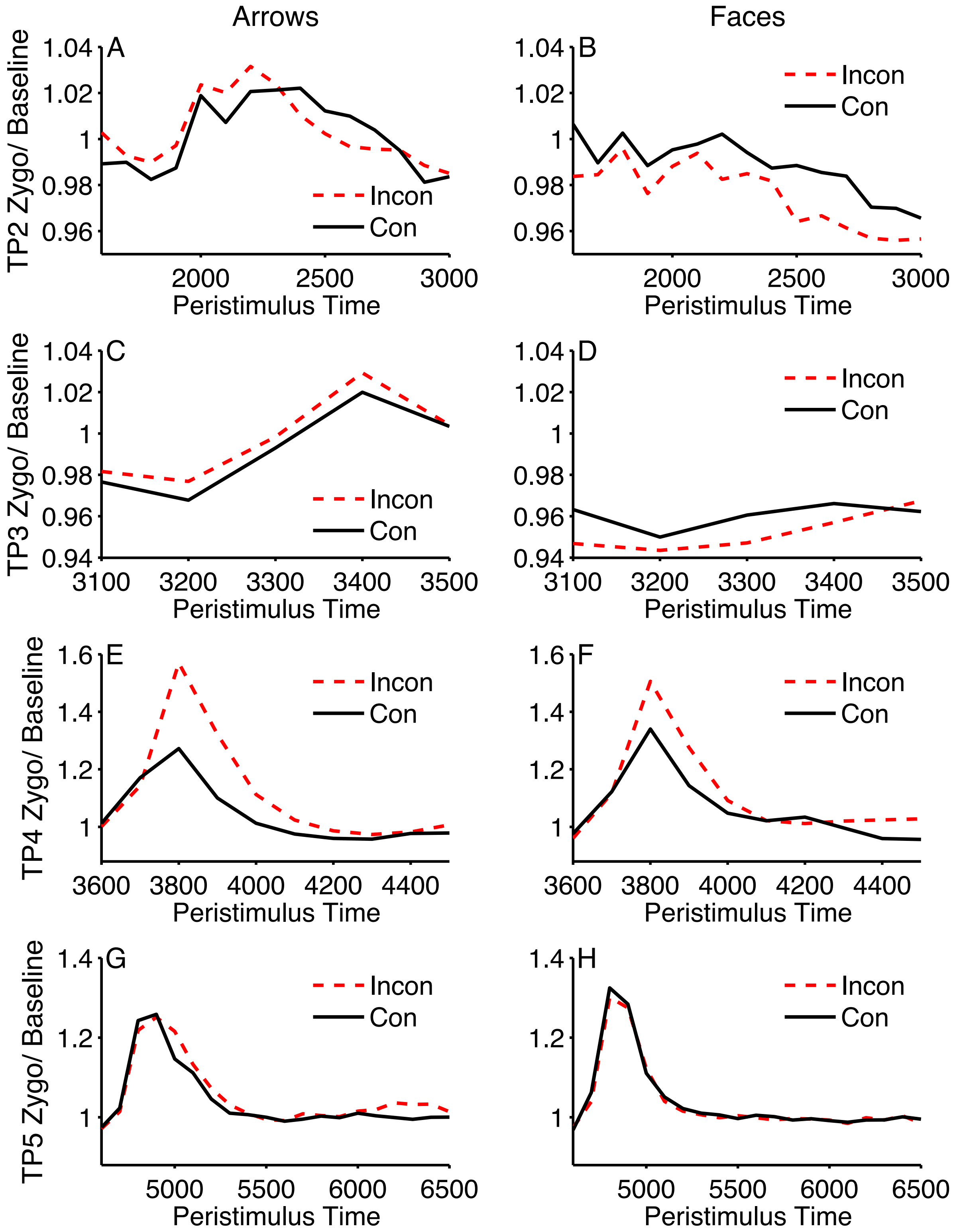


*Figure 37.* Mean stimulus-locked zygomaticus activity on congruent (solid line) and incongruent trials (dashed line) for arrows (left panels) and faces (right panels) across trial periods 2, 3, 4 and 5 (rows). EMG units on the Y-axis represent the ratio of activity relative to baseline (fixation).

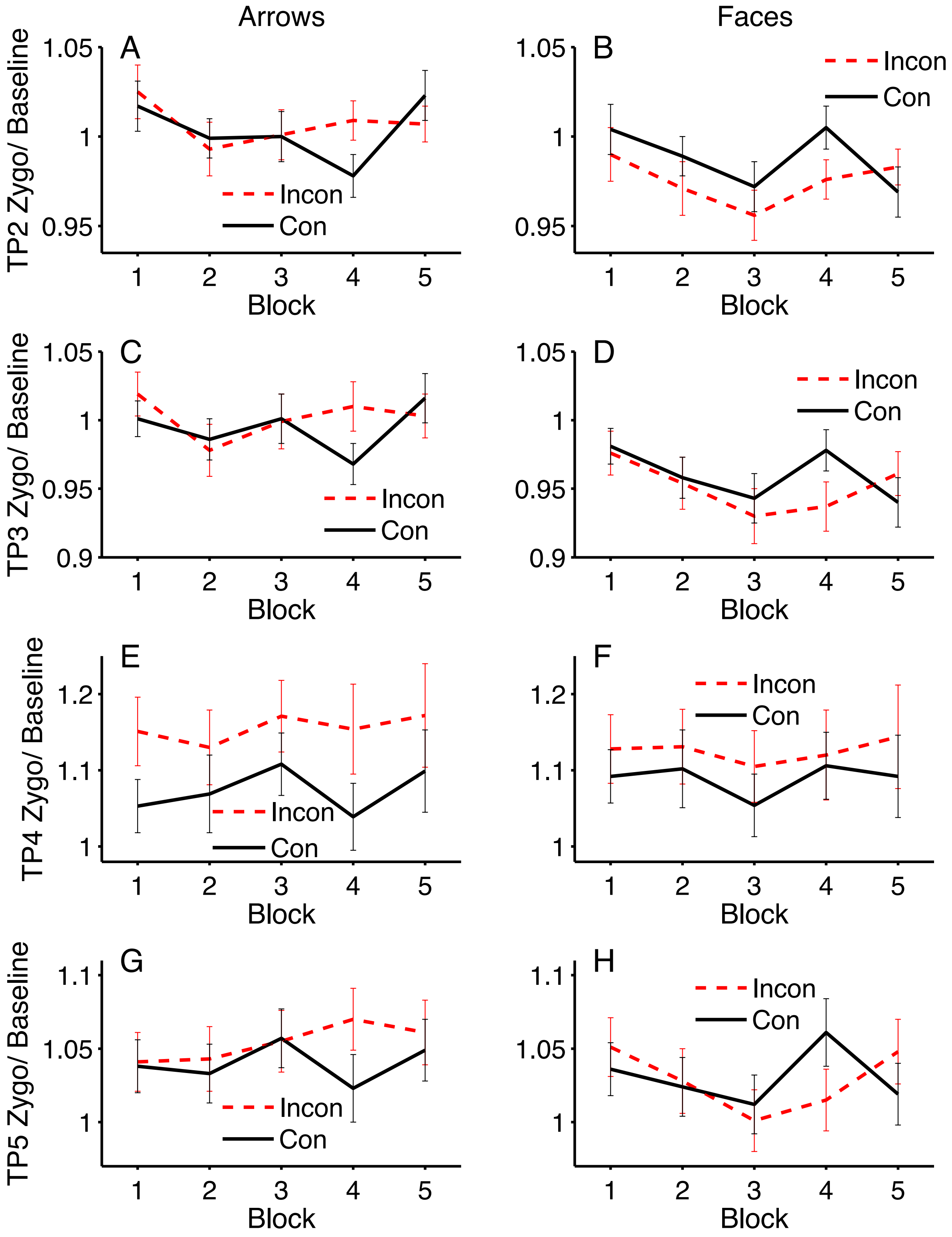


*Figure 38.* Mean zygomaticus activity across blocks for arrows (left panels) and faces (right panels) in trial periods 2, 3, 4 and 5 (rows). Dashed lines represent incongruent and solid lines represent congruent. EMG units on the Y-axis represent the ratio of activity relative to baseline (fixation). Error bars show +/- 1 standard error of the mean.

## Discussion

Experiment 5 aimed to investigate whether the propensity of emotional responses to gaze-cues to elicit changes in evaluations was specific to the social nature of faces and gaze or whether it is due to the domain general properties of attentional cueing. To this end, we compared the learning of liking of faces from gaze-cueing to the learning of liking of arrows from arrow-cueing and examined whether emotion may mediate this learning by measuring facial EMG online. Despite large differences in the visual attributes of the arrows and faces, both types of stimuli were equivalent in terms of perceptual discriminability, memorability and attentional cueing as measured in the visual search, memory and cueing experiments.

Yet only faces elicited changes in liking that differed between congruent and incongruent attentional cues. Unlike previous experiments, this was driven solely by the congruent faces increasing in liking from before to after cueing. This liking effect was also much weaker than that previously observed for trust suggesting that trust may be a more pertinent evaluative descriptor for faces and gaze cues (Oosterhof & Todorov, 2008). While the emotional reactions and learning of trust may have been similar to previous experiments, the liking scale may be used in a different manner to trust. Liking may be more of a unipolar concept compared to trust, ranging from no liking to extreme liking- more of a focus on how positive a stimulus is. Participants may have used the scale in a corresponding manner. Nevertheless, the weaker effects of identity-contingent gaze-cues on liking compared to trust is consistent with previous studies showing that differences in preference judgements were weaker and less consistent than trust (Bayliss et al., 2013; Bayliss & Tipper, 2006).

There was no significant effect of validity on the liking of arrows despite an overall negative change in ratings. This is consistent with previous studies showing object liking is increased when consistently cued towards than away by gaze but not arrows or pointing hand cues (Bayliss et al., 2006; 2013; Capozzi et al., 2014; Manera et al., 2014; Treinen et al., 2012; Ulloa et al., 2015). However, there was a slight difference in the effect of validity across blocks and stimuli. Whereas the effect of validity was constant throughout all blocks for faces/ gaze, the arrow cueing effect was diminished in block 5. It is generally accepted that gaze and arrows cue attention in a very similar manner (Bayliss & Tipper, 2005; Tipples, 2002; 2008). Our data may suggest that differences in the ability of arrows and gaze to cue attention may arise after participants learn that particular arrows/ faces consistently cue congruently or incongruently. Indeed, Tummeltshammer et al. (2014) found that infants overtly followed the gaze of congruent faces but not incongruent faces, where as there was no difference between incongruent and congruent arrows. This may be related to the knowledge that eye-movements can be under voluntary control, depend on mental state and provide more valuable information about mental states than arrows (Teufel et al., 2010).

The EMG data provide complementary evidence to the ratings. The larger EMG activity to arrows compared to faces observed in the corrugator in trial period 2 and in the zygomaticus in trial periods 2 and 3 is consistent with the generally more negative ratings given to arrows at the end of the experiment. Perhaps this greater negative emotional response to arrows before the gaze-cue prohibits interest or motivation to learn the valence of the outcome of congruent/ incongruent attentional cues in trial period 4. Indeed, in natural scenes containing arrows and faces, participants fixate faces and eyes first and more often than arrows suggesting they are more interesting (Birmingham et al., 2009). As in previous experiments, a larger EMG reaction was found in both the corrugator and zygomaticus to incongruent compared to congruent cueing in trial period 4. This effect was obtained for both arrows and faces. This suggests that the emotional response to faces seen in trial period 4 in this and other experiments is a property of domain-general attentional cueing that can be elicited by non-social cueing stimuli such as arrows. However, for arrows, this does not appear to relate strongly to the learning of liking as there was no overall liking effect in this group. In comparison, overall there was a liking effect for faces. This difference could be due to the extent and duration of the emotional reaction. The larger emotional response to incongruent compared to congruent was sustained from trial period 4 into trial period 5 for faces but not arrows. The fact that this interaction is observed in a trial period in which explicit congruency information is not visible suggests that the emotional reaction is more likely to reflect the unique combination of the face with learned emotion from congruency, two pieces of information that are more likely to be associated and remain in memory. The effect in trial period 5 could be understood within the iterative reprocessing model of evaluation according to which reprocessing of initial automatic implicit evaluations leads to more reflective explicit judgements that integrate contextual information, such as face identity, into the evaluation (Cunningham & Zelazo, 2007). The relatively late effect in trial period 5 may be due to the accrual of a sufficient number of evaluative iterations.

There was also an interaction between validity, stimulus and block in the zygomaticus in trial periods 3 and 5. However, the validity × block interactions did not reach significance when analysed in each group which questions whether the effect contributes anything important to learning. At a finer-grained level, there was no interaction between liking effect and EMG activity for faces. This is in contrast to experiment 2 where the difference in corrugator activity in trial period 4 was larger in trust effect participants. There could be several reasons for this. The relation between trust effect and validity in experiment 2 was a weak effect, only just passing the threshold for significance. Combined with the weaker overall effect in the means and the possibility that the concept of liking may not relate to face evaluation as closely as trust and may lead to more of a focus on the positive aspects of stimuli present in congruent faces, EMG activity may not so closely relate to liking ratings.

In conclusion, experiment 5 has investigated whether the learning of evaluations, and a mediating role for emotion, is specific to the social qualities of faces and gaze or is due to domain-general attentional orienting which can also be elicited by arrows. The findings suggest that while the congruency of arrows can elicit a short-lived emotional response in trial period 4 when congruency is explicit, it is more sustained in response to faces as it remains in trial period 5. This may be why learning of liking based on validity is only observed for faces coupled with the fact that arrows are less interesting and participants have less incentive motivation to learn about these stimuli. Although a liking effect was observed for faces, the results support the notion of a special relationship between faces, gaze, emotion and trust, as the liking effect was weaker than the trust effects previously observed.

# Chapter 7

## Experiment 6: An investigation of whether the learning of liking from gaze-cueing is reliable and underpinned by the same neural mechanisms as the learning of trust

Experiment 5 found an effect of identity-contingent gaze-cueing on the learning of liking judgements of faces. However, this effect was much weaker than the effects obtained in experiments 1, 2, 3 and 4 measuring trust judgements despite equivalent embodied emotional reactions as measured with EMG. The aims of experiment 6 were to attempt to replicate and clarify the findings observed in previous experiments. Specifically, we were interested to see whether the liking effect observed in experiment 5 could be replicated and whether it is related to similar learning effects in the LPP ERP as seen in experiment 1. The LPP is modulated by cognition as evidenced by its down- or up-regulation by neutral or negative descriptions of observed affective pictures (MacNamara, Foti & Hajcak, 2009), attention orienting to non-arousing portions of the images (Hajcak et al., 2009) or through cognitive reappraisal to think about pictures in a more positive/ negative way (Hajcak & Nieuwenhuis, 2006; Moran et al., 2013). In experiment 5 we found that the learning of liking was mainly driven by a positive change in ratings to congruent faces. This was interpreted as being due to the nature of the concept of liking which may engender an evaluation that is more focused on a stimulus's positive attributes than its negative attributes. Given this sensitivity of the effect of gaze-cues on evaluations to the way faces are framed or described, perhaps one would also expect corresponding changes in its neural representation.

In experiment 4 we found that there was no interaction between appearance-based trustworthiness and the learning of trust from gaze-cues. This could possibly have been due to the initial rating of faces. Making decisions about trustworthiness in the initial rating based on facial appearance may have disrupted finer-grained learning of trust from gaze as suggested by the findings of Delgado et al. (2005). Perhaps initial liking judgements also bias the way in which identity-contingent gaze-cues are processed and evaluated. In this view, it is not just that liking ratings don't relate to faces as much as trust, but that initial liking ratings also bias the way in which the faces are encoded. A previous experiment has shown that the LPP to positive stimuli is larger compared to negative and neutral stimuli when stimuli are classified as positive or not positive whereas the LPP is larger to negative compared to positive and neutral stimuli when stimuli are classified as negative or not negative (Rehmert & Kisley, 2013). Therefore, given the plasticity of the LPP depending on framing, and the outcome of experiment 5, showing that liking may engender a focus on the positive aspects of faces, as accessible from the congruent faces, we anticipated that the effect of gaze-cues on the LPP would be reversed or at least different to that observed with trust in experiment 1. This may or may not depend upon a change in the liking ratings in response to gaze-cueing, as, given the weaker effect for liking and the inconsistency across studies measuring preference, the liking effect observed in experiment 5 may not be replicated. If LPP effects are detected in spite of no effects of liking, this may further suggest dissociation between this measure of implicit emotional responses and explicit evaluations as observed in experiment 1 where there were no interactions of the LPP with trust effect. We also used a different set of faces to all previous experiments without the pupil size manipulation as in experiment 1.

## Method

### Participants

There were 22 participants in total, of which 21 were right handed and of which 16 were female. The mean age of participants was 20 (*SD* = 1.6). All participants were neurologically normal with normal or corrected to normal vision. All procedures were approved by the Bangor University ethics committee.

### Stimuli and Design

Eight male and 8 female face stimuli were taken from the Glasgow unfamiliar face database (Burton, White & McNeill, 2010; Faces 021, 022, 032, 054, 071, 097, 128, 174, 175, 185, 194, 208 222 and 223) and the Utrecht face database (pics.psych.stir.ac.uk; Faces M4077 and M4012_1). The faces were selected and grouped into A and B groups based on equivalence in happiness, trustworthiness and liking judgements given by a group of 16 independent observers who had a mean age of 21 (*SD* = 5) and were mostly female (*N* = 14) and right handed (*N* = 15). Happiness (*M* = 77.5, *SD* = 11.8) was rated from neutral (0) to very happy (+100), trust (*M* = 8.9, *SD* = 9.6) was rated on a scale from very untrustworthy (-100) to very trustworthy (+100), and liking (*M* = .6, *SD* = 14.0) was rated from dislike (-100) to like (+100). There were no significant differences between counterbalanced face groups in terms of happiness, *t* (15) = -.253, *p* = .804, 95% CIs [-3.8, 2.96], trust, *t* (15) = 1.7, *p* = .107, 95% CIs [-1.16, 10.6], or liking, *t* (15) = .482, *p* = .637, 95% CIs [-9.8, 15.6], in these preliminary ratings. Unlike in experiment 1, the pupils of the faces were not artificially adjusted to be larger for congruent faces and smaller for incongruent faces. Although the faces were not matched for similarity, the counterbalancing of faces designated congruent or incongruent across participants and taking of change scores ensured that initial differences in liking were minimal and controlled for in the end ratings. In all phases of the experiment, the faces were at a size of 527 × 685 pixels and the objects were 350 × 263 pixels. The experiment was run using E-Prime 1.0 (Psychology Software Tools, Inc.) with a 1920 × 1080 screen resolution on a 24” Samsung SyncMaster BX2431 LED display, which was 569 × 342 mm in dimensions and had a 500 Hz refresh rate. In all other respects, the experiment was the same as experiment 1, apart from liking ratings were measured instead of trust.

### EEG Recording and Pre-processing

The EEG recording and pre-processing was exactly the same as experiment 1 apart from trials with reaction times below 500 ms were removed from analyses of all trial periods and the size of the images was 32 × 32 mm across all time points for viewing trials.

### Data Screening Protocol

Errors, artefacts and reaction time outliers above 1500 ms or below 500 ms were removed from all of the EEG analyses. Using paired samples t-tests, there was no significant difference between incongruent and congruent conditions in terms of the number of trials removed due to errors, $t$ (21) = .099, $p$ = .922, 95% CIs [-.57, .62], outliers, $t$ (21) = -.37, $p$ = .715, 95% CIs [-.75, .52], or artefacts, $t$ (21) = 1.104, $p$ = .282, 95% CIs [-.48, 1.6]. For the reaction time data, errors were removed along with trials exceeding two standard deviations above or below each participants mean reaction time or a maximum of 1500 ms. There was a significant difference in the number of outliers removed from the reaction time data $t$ (21) = 2.2, $p$ = .041, 95% CIs [-.71, .23], in the expected direction, with there being more outliers in the incongruent condition ($M$ = 2.93%, $SD$ = 1.5%) compared to the congruent condition ($M$ = 2.13%, $SD$ = 1.5%) (see Table 11). A 2 × 5 repeated measures ANOVA with factors of validity and block was run on the number of trials in the cueing EEG analysis (see Table 12). There was no significant effect of validity, $F$ (1,21) = .313, $p$ = .582, $\eta_p^2$ = .015, block, $F$ (4,84) = .611, $p$ = .656, $\eta_p^2$ = .028, or a block × validity interaction, $F$ (4,84) = .393, $p$ = .813, $\eta_p^2$ = .018. Another 2 × 2 ANOVA was run on the number of viewing trials, with factors of time and validity. There was no significant effect of time, $F$ (1, 21) = 0.0, $p$ = 1.0, $\eta_p^2$ = 0.0, validity, $F$ (1, 21) = .082, $p$ = .777, $\eta_p^2$ = .004, or an interaction, $F$ (1, 21) = .307, $p$ = .585, $\eta_p^2$ = .014, before artefacts were removed, and no significant effect of time, $F$ (1, 21) = .1.81, $p$ = .192, $\eta_p^2$ = .08, validity, $F$ (1, 21) = 2.39, $p$ = .137, $\eta_p^2$ = .102, or an interaction, $F$ (1, 21) =

.213, $p = .649$, $\eta_p^2 = .01$, after artefact removal (see Table 13). As in previous experiments, participants were divided into those who showed liking effects (10 participants) and those who did not show a liking effect (12 participants).

Table 11. *Means and standard deviations of percentages of errors, EEG outliers, reaction time outliers and artefacts removed from each condition in the cueing analyses.*

| | Errors | EEG outliers | Reaction time outliers | Artefacts |
|---|---|---|---|---|
| Congruent | 1.9(1.3) | 2.1(1.7) | 2.1(1.5) | 4.9(4.3) |
| Incongruent | 1.9(1.6) | 2.8(1.7) | 2.9(1.5) | 5.5(3.9) |

Table 12. *Means and standard deviations of the numbers of trials in each condition of the cueing analyses.*

| | Block 1 | Block 2 | Block 3 | Block 4 | Block 5 |
|---|---|---|---|---|---|
| Congruent | 13.1(2.2) | 13.3(2.16) | 13.1(2.4) | 13.6(2.1) | 12.7(2.4) |
| Incongruent | 13.5(2.7) | 13.4(1.7) | 13.4(1.9) | 13.1(2.9) | 12.8(2.6) |

Table 13. *Means and standard deviations of the numbers of repetitions and number of trials remaining in the viewing analysis conditions after artefact rejection.*

| | Beginning | | End | |
|---|---|---|---|---|
| | Congruent | Incongruent | Congruent | Incongruent |
| Repetitions Mean | 5.9(.35) | 6.1(.35) | 6.0(.54) | 6.0(.54) |
| Number of Trials Mean | 45.6(4.2) | 44.5(4.6) | 45.1(4.43) | 42.9(6.2) |

## Results

### Liking Ratings

The liking ratings were analysed with a 2 × 2 × 2 within-subjects ANOVA with factors of time of rating, face gender and validity. There were no significant effects of validity, $F(1, 21) = .095$, $p = .761$, $\eta_p^2 = .005$, or time of rating, $F(1, 21) = 1.43$, $p = .245$, $\eta_p^2$

= .064, nor was there an interaction between time of rating and validity, *F* (1, 21) = .172, *p* = .682, $\eta_p^2$ = .008. However, there was a significant main effect of face gender, *F* (1, 21) = 12.24, *p* = .002, $\eta_p^2$ = .368, due to higher liking ratings for females (*M* = 12.0, *SEM* = 4.06) compared to males (*M* = -2.6, *SEM* = 4.7) (see Figure 39). No other interactions reached significance.

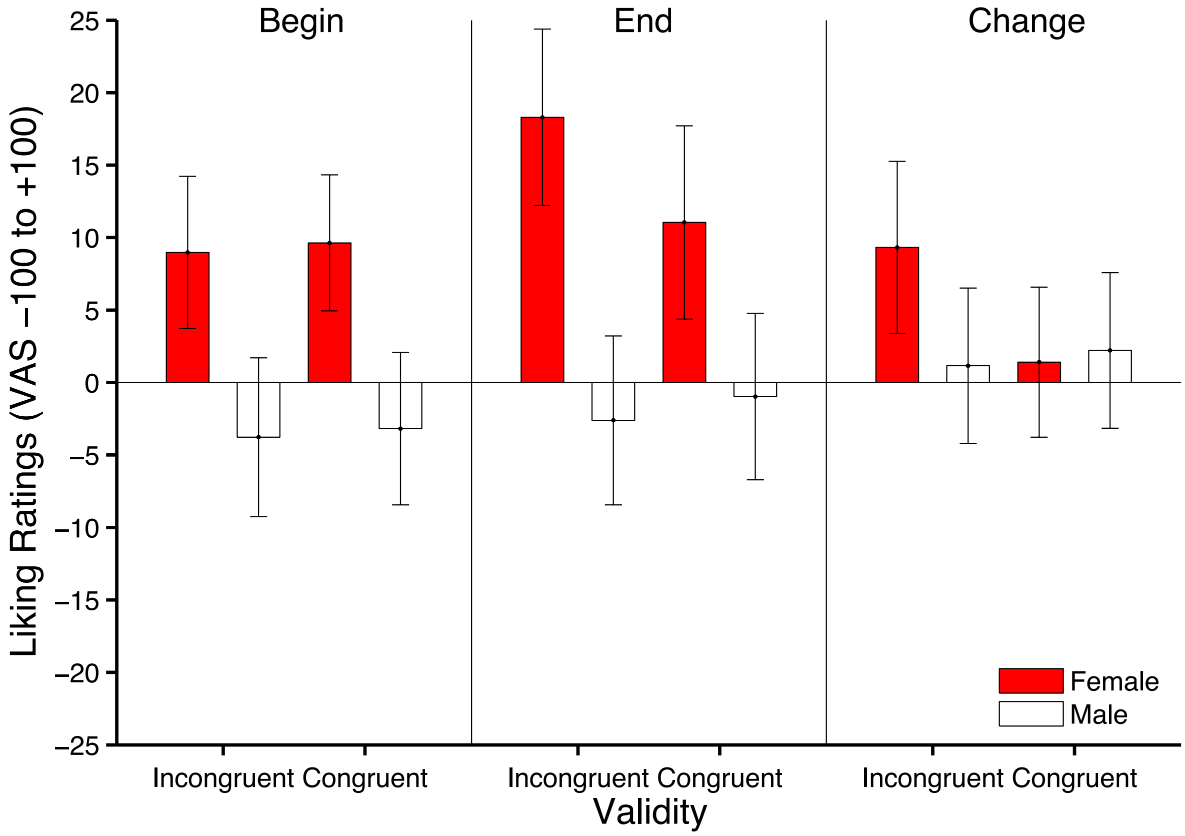


*Figure 39.* Liking ratings before (left panel) and after cueing (middle panel) and the change in liking ratings from before to after cueing (right panel) (end-beginning). Error bars represent +/- 1 standard error of the mean.

## Gaze-Cueing Reaction Times

The cueing data were analysed with a 2 × 2 × 2 × 5 mixed ANOVA with the between subjects factor of liking effect and the within-subjects factors of face gender, validity and block. There was a significant main effect of validity, *F* (1, 20) = 12.76, *p* = .002, $\eta_p^2$ = .389, owing to faster reaction times on congruent trials (*M* = 810.4, *SEM* = 25.9) compared to incongruent trials (*M* = 836.9, *SEM* = 23.3) (see Figure 40). An interaction between liking effect and validity approached significance, *F* (1, 20) = 3.81, *p* = .065, $\eta_p^2$ = .16, due to a larger gaze-cueing effect for the liking effect participants compared to the no liking effect participants. There was also a significant effect of block, *F* (4, 80) = 25.28, *p* < .0001, $\eta_p^2$ =

.56, due to a linear decrease in reaction times as blocks progressed *F* (1, 20) = 49.8, $p <$ .0001, $\eta_p^2 = .714$. Given the borderline interaction between validity and liking effect, subsequent 2 × 2 × 5 within-subjects ANOVAs were run in the liking and no liking groups. This analysis showed that the effect of validity was significant in the liking effect group, *F* (1, 9) = 9.12, $p = .014$, $\eta_p^2 = .50$, but not in the no liking effect group, *F* (1, 11) = 2.56, $p = .138$, $\eta_p^2 = .189$.

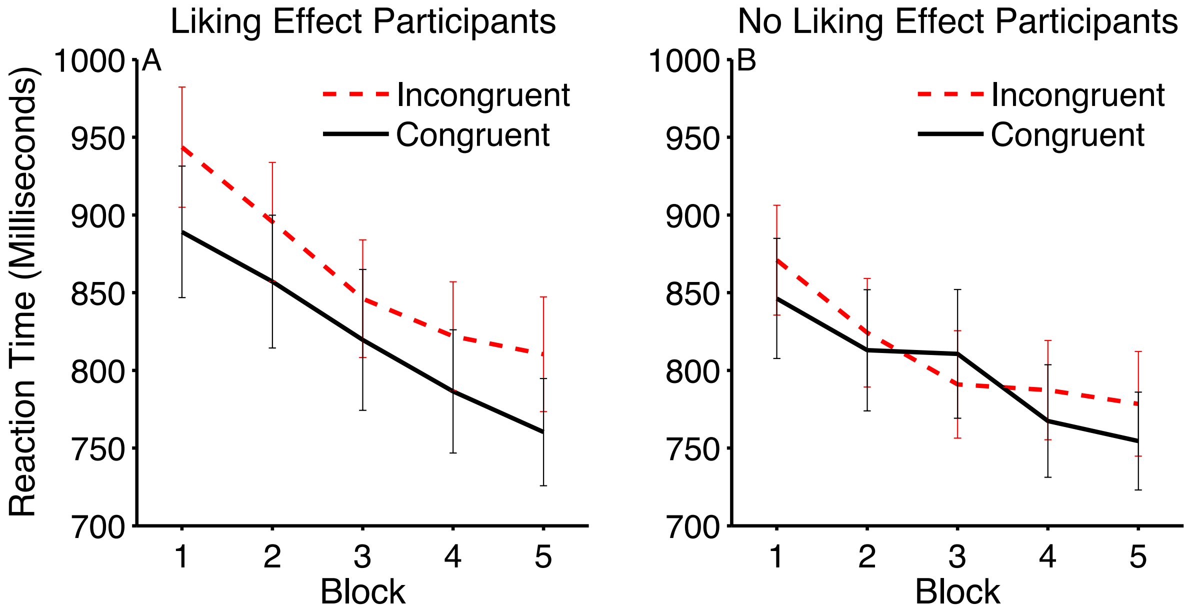


*Figure 40*. Gaze-cueing reaction times across blocks for liking effect participants (left panel) and no liking effect participants (right panel). Solid lines represent congruent and dashed lines represent incongruent. Error bars represent +/- 1 standard error of the mean.

**Passive Viewing Phase ERPs**

Averaged scalp-time images were analysed with a factorial ANOVA with factors of time and validity. All effects were thresholded at $F(1,80) = 6.96$, $p < .01$, uncorrected. There was a significant main effect of time as evidenced by two separate clusters of activity on central ($p = .030$, $k = 16969$) and posterior occipito-temporal ($p = .040$, $k = 15662$) electrodes between 170-288 ms, peaking at 219 ms, $F(1,80) = 19.74$, $p < .001$, 217 ms, $F(1,80) = 16.3$, $p < .001$, and 260 ms, $F(1,80) = 12.4$, $p < .01$, for the former and at 186 ms, $F(1,80) = 12.3$, $p < .01$, 211 ms, $F(1,80) = 10.6$, $p < .01$, and 217 ms, $F(1,80) = 9.33$, $p < .01$, for the latter. The peaks in the posterior cluster were on the left hemisphere at 186 ms and on the right hemisphere at 211 and 217 ms. Figure 41 shows the scalp distribution of the effect and the activity on peak electrodes in the anterior (D14) and posterior (B12) clusters at 219 and 211 ms, respectively. The effect reflects the N250 component with the signal being larger at the beginning compared to the end on posterior occipito-temporal electrodes and larger to end compared to beginning on central electrodes at around 200-300 ms post-stimulus (see Figure 41). No other main effects or interactions, including those with liking effect, were significant. These findings are consistent with experiment 1, which also showed main effect of time at roughly the same latency and scalp distribution and no effects or interactions with validity.

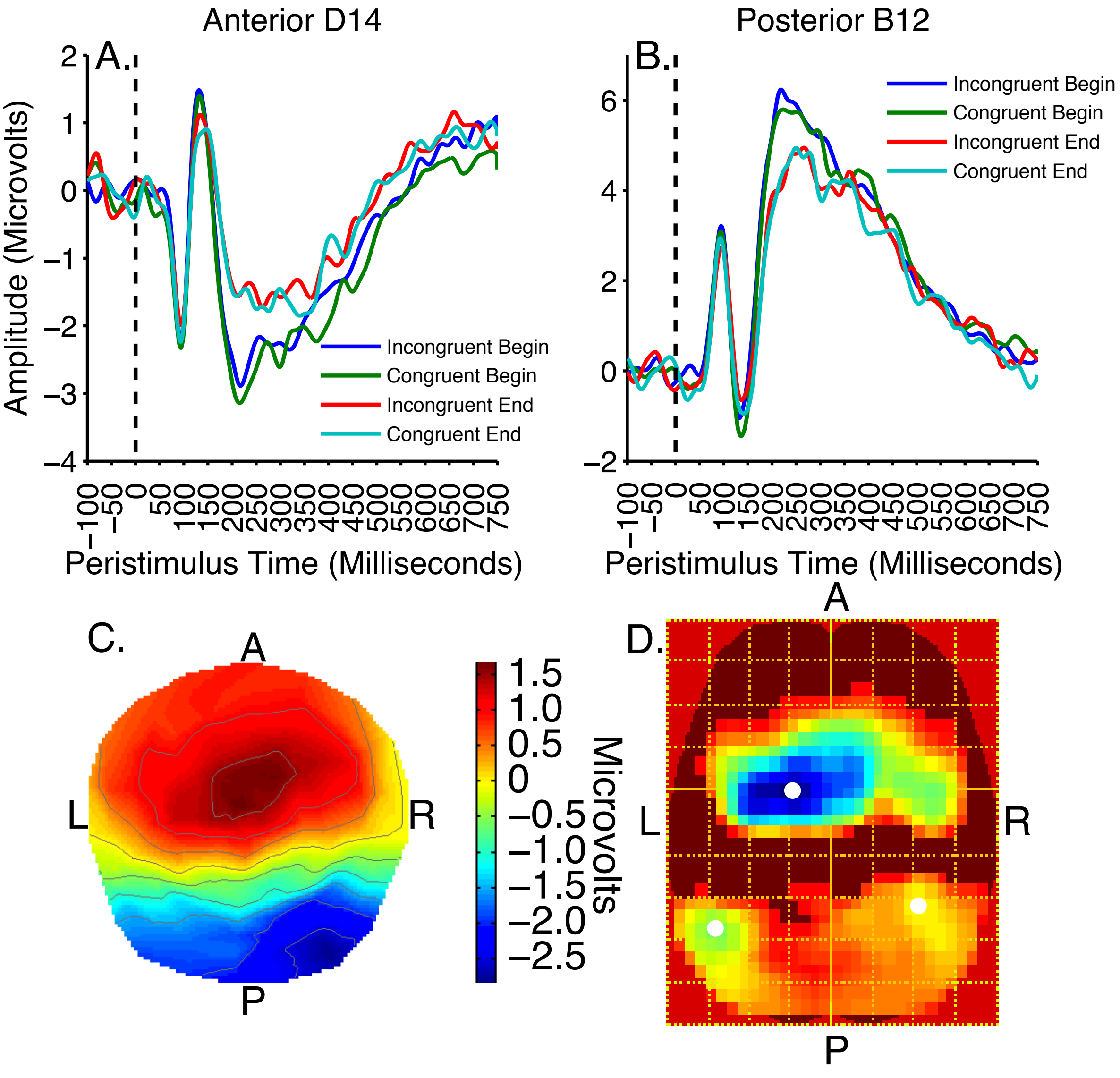
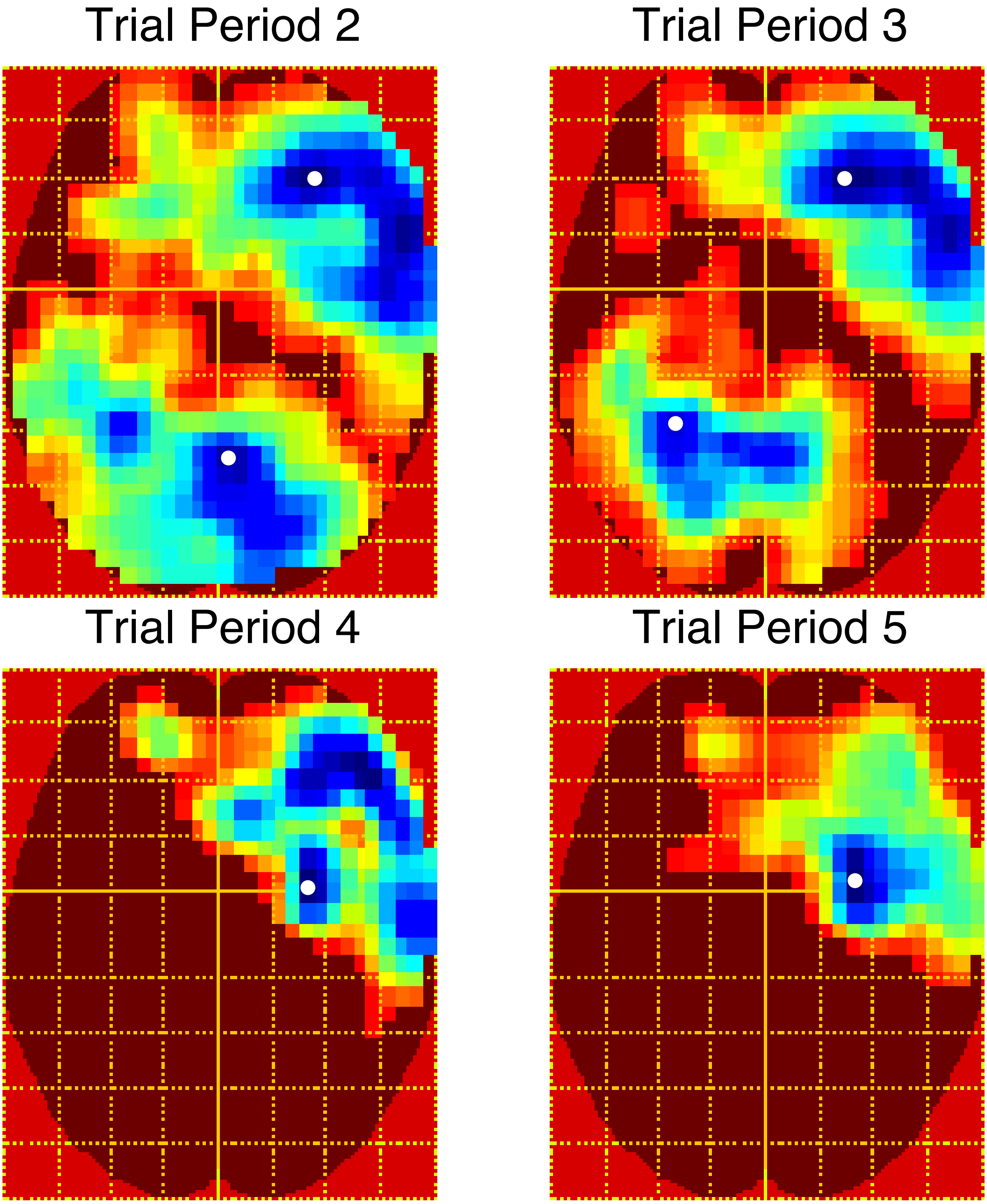


*Figure 41*. ERP activity on viewing trials, showing N250 effects. Top panels (A) and (B) show activity on the anterior (D14) and posterior (B12) sensors closest to where the N250 effect peaked in each cluster. Bottom panels (C) and (D) show the scalp distribution of the N250 effect at 219 ms and the F-map of significant voxels on the scalp throughout the trial. White spots show peaks.

### Cueing Phase ERPs

All contrasts were submitted to a cluster-forming height threshold of $F(1,420) = 3.86$, $p < .05$, uncorrected. Clusters were only considered significant if they passed a clustersize threshold of at least $p < .05$, corrected for FWE with RFT. Also, the block × validity interaction was investigated with the same contrast weights used in experiment 1, with congruent faces weighted as 2, 1, 0, -1, and -2 across blocks and with the sign of incongruent weights being reversed. There was a significant effect of validity in trial periods 2 and 3 as evidenced by two large clusters of activity. The LPP was larger to incongruent compared to congruent faces on right central-anterior sensors whereas the polarity was reversed on posterior occipito-temporal sensors. In trial period 2, the effect in both clusters was significant continuously from around 250 ms until the end of trial period 2, peaking at 1359 ms (on C10) for the anterior cluster and at 842 ms (on A20) for the posterior cluster. This main effect of validity extended into trial period 3, sustaining the same scalp distribution as in trial period 2. Both the anterior and posterior clusters were significant throughout the whole trial period. The peak of the anterior cluster occurred at 74 ms (on C10) and the peak of the posterior cluster occurred at 75 ms (on A7). This effect of validity only continued to remain significant throughout trial period 4 and 5 on anterior electrodes nearest B31, peaking at 310 ms for trial period 4 and 746 ms in trial period 5 (see Figures 42 and 43). The main effects of validity in trial periods 2 and 3 were still highly significant when only blocks 2 to 5 were included in the contrast and did not differ to the contrast with block 1 included.

In trial period 2, there was also a significant linear block × validity interaction on mid-central sensors between 250-1270 ms ($p < .0001$, $k = 95681$, RFT FWE corrected) and posterior occipital sensors between 734-1290 ms ($p = .022$, $k = 50035$, RFT FWE corrected) peaking at 836 ms (on C2) for the former and at 1116 ms (on B8) for the latter (see Figure 44). The effect appears to be due to a larger difference between incongruent and congruent faces in block 5 compared to earlier blocks, this difference being positive on central sensors and negative on posterior sensors (see Figure 45). In order to test whether the linear block × validity interaction was driven by congruent or incongruent faces, the effect of block was tested separately using the contrast weights -2, -1, 0, 1 and 2 across blocks 1 to 5. This analysis revealed that the anterior and posterior clusters remained significant for congruent but not incongruent faces. All main effects and interactions were tested with liking effect as a between-subjects covariate. However, no interactions reached significance.



Trial Period 2

Trial Period 3

Trial Period 4

Trial Period 5

*Figure 42*. *F*-maps showing significant effects of validity on the scalp across trial periods (voxelwise thresholded at $F$ (1, 420) = 3.86, $p < .05$; clustersize significant at $p < .05$ RFT FWE corrected). White spots denote location of peak effects.

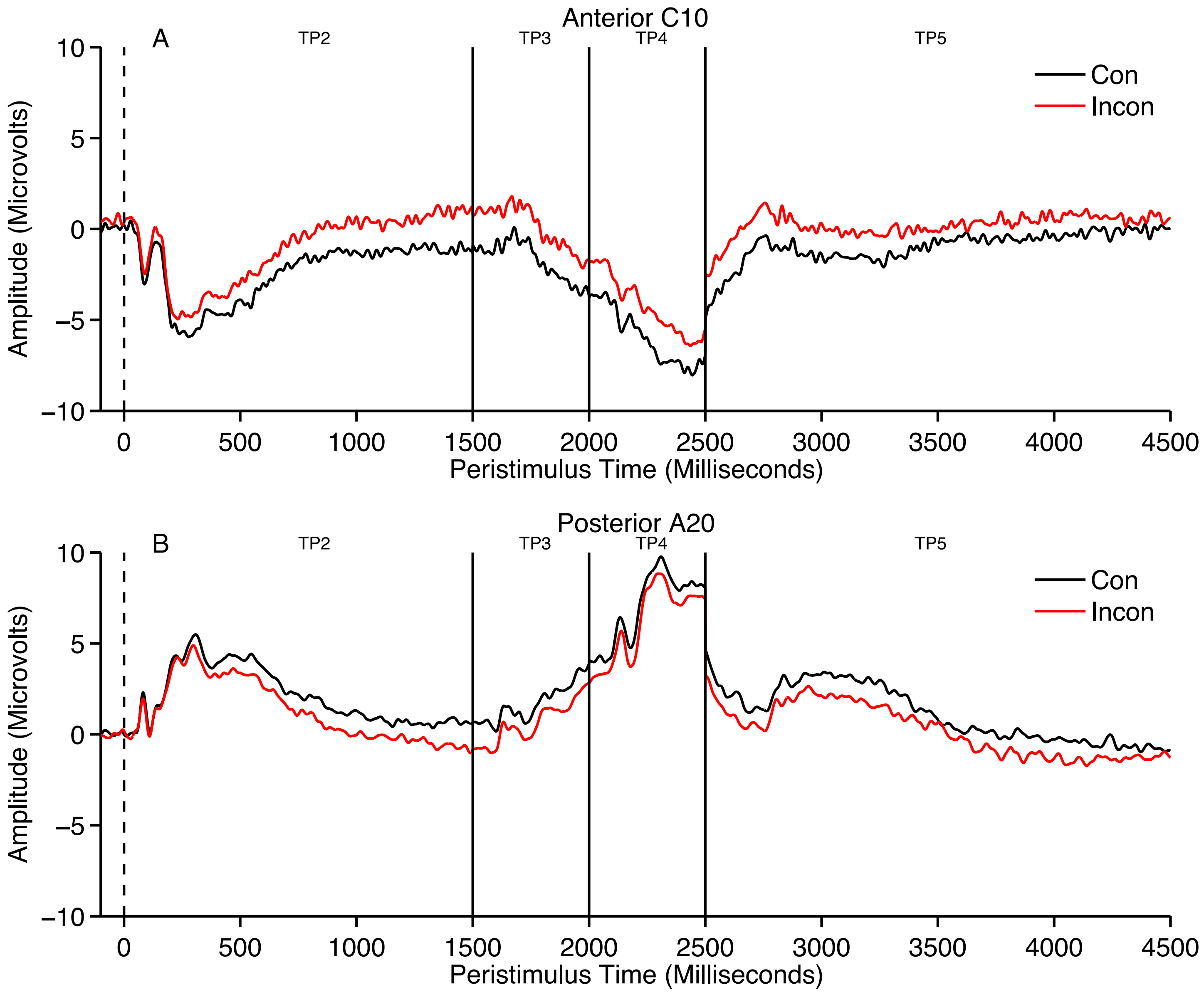


*Figure 43*. Time course of activity on incongruent (red line) and congruent trials (black line) on right frontal electrode C10 (top panel) and midline parietal electrode A20 (bottom panel) located in the anterior and posterior clusters, respectively.

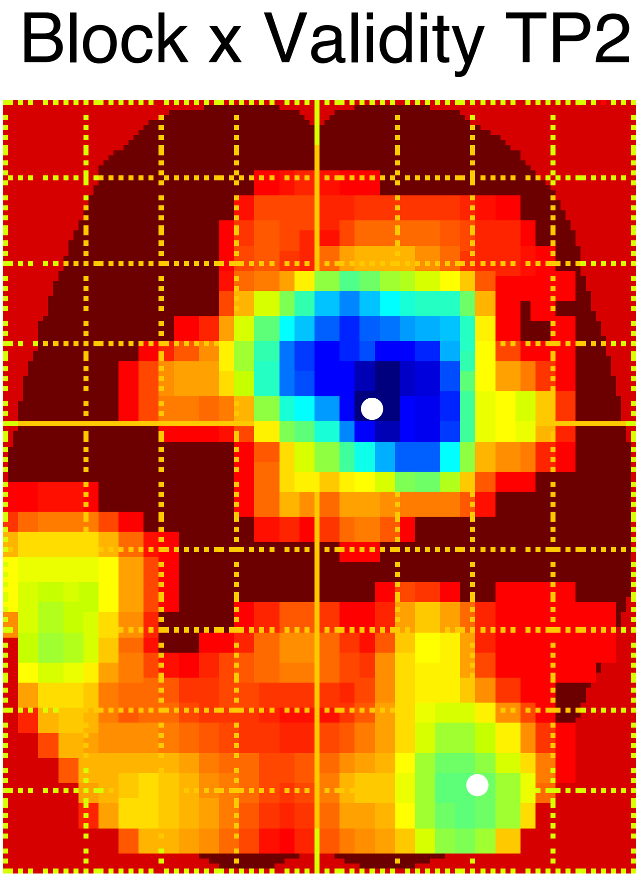


*Figure 44*. *F*-map showing the significant linear block × validity interaction on the scalp in trial period 2 (voxelwise thresholded at $F\ (1, 420) = 3.86$, $p < .05$; clustersize significant at $p < .05$ RFT FWE corrected). White spots denote location of peaks.

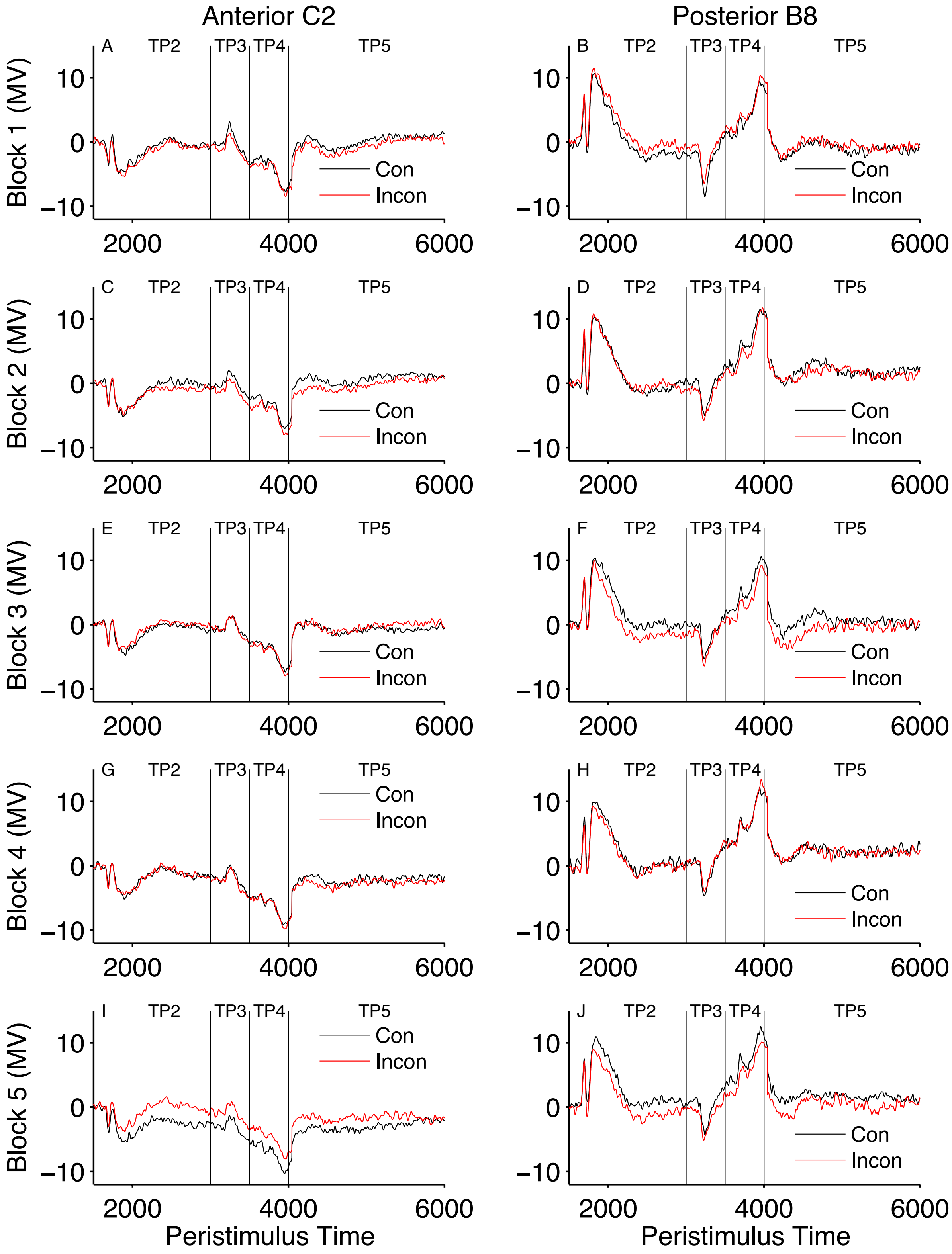


*Figure 45.* ERP activity on incongruent (red lines) and congruent (black lines) trials across blocks (rows) on peak electrodes in the anterior cluster (C2) and the posterior cluster (B8).

## Discussion

Experiment 6 aimed to address outstanding questions from previous experiments. We were interested to see if the effect of identity-contingent gaze-cues on liking judgements was reliable and if similar ERP responses are elicited compared to when trustworthiness is assessed which might suggest a potential influence of the initial ratings on learning. We were also interested to see whether the effect generalised to a new stimulus set, the eyes of which were not manipulated to have small or large pupils as in experiment 1. In line with the weaker effects of gaze on liking ratings in experiment 5 compared to trust ratings in all other experiments, we found that there was no effect of gaze on liking suggesting that liking does not relate to faces and gaze as readily as trust. This is consistent with studies showing that preference judgements of faces are less reliably changed by identity-contingent gaze-cues than trust (Bayliss et al., 2013; Bayliss & Tipper, 2006). Despite this, we detected significant effects of validity on the LPP throughout the task, although these were weaker than that observed in experiment 1. The LPP was larger to congruent faces on posterior electrodes. This was coupled with a reversal of the effect on right anterior sensors, which is consistent with previous studies showing a larger right lateralized LPP to concepts rated as bad compared to good (Cunningham et al., 2005) and with a long line of evidence for right prefrontal involvement in negative affect (Harmon-Jones et al., 2010).

However, it is unclear whether the effect is due to a larger positive emotion to congruent faces, as suggested by the effect in the posterior cluster, or due to a larger negative response to incongruent as suggested by the right anterior cluster. Whilst the LPP is typically measured as a positivity from posterior parietal electrodes (Schupp et al., 2000), there is strong evidence for right prefrontal involvement in negative affect (Harmon-Jones et al., 2010). As the effect was stronger on the right anterior electrodes, we suggest that the latter is more likely to be the case. This would be consistent with experiment 5 where corrugator activity was larger for incongruent faces in trial periods 4 and 5. While the findings suggest that the negative, untrustworthy, incongruent gaze-cues may elicit a larger emotional response than trustworthy congruent faces, the spatial distribution and polarity of the effects were much different to experiment 1. Therefore there may well be subtle differences in the way identity-contingent gaze-cues are processed depending on whether participants are primed with liking or trustworthiness in the initial ratings. The LPP is not reflective of an isolated function and may recruit different sets cognitive processes that are know to interact with emotion such as perception, attention and memory depending on priming.

The significant ERP findings and the non-significant effects in the liking ratings would suggest dissociation between implicit emotional responses measured via the LPP ERP and explicit liking judgements which would also be consistent with experiment 5 where there was no fine-grained relationship between liking ratings and EMG activity. Again, this may be explained in terms of a focus on how positive a stimulus is regardless of its negative attributes. Indeed, there was a general non-significant increase in liking ratings from before to after cueing, which could possibly also reflect a mere exposure effect, typically measured with liking (Zajonc, 1968). Although there appeared to be an interaction between block and validity in trial period 2, this was much weaker than in experiment 1 and was confined only to this trial period. This may also explain why no significant learning effects were observed in the liking ratings. It was also shifted more posteriorly, was in the reverse direction and driven primarily by block 5, suggesting that it took longer to learn face-gaze contingencies. It was driven primarily by congruent faces which is in agreement with experiment 5 where only the positive change to congruent faces was responsible for the weaker liking effect. The fact that effects are weaker for the positive congruent gaze-cues and liking ratings that might engender a focus on positivity is testament to the strength of the negativity bias.

Although the ERP effects did not appear to relate to liking, participants who showed a liking effect also showed a greater effect of validity in the reaction times compared to participants who did not show a liking effect. This is in contrast to experiments 1, 2, 3 and 4 where there was no interaction between trust effect and validity in the trust ratings. This might suggest that, unlike trust ratings, effects in the liking ratings are driven primarily by fluency. Effects of fluency are most commonly observed in liking ratings (Cannon et al., 2010; Regenberg et al., 2012; Winkielman & Cacioppo, 2001; Winkielman et al., 2006) and could be related to the ERP effects. A distinction has previously been made between liking and wanting, the former of which is related to the experience of pleasure and the latter to motivation to obtain pleasure (Berridge & Robinson, 1998). Schilbach et al. (2010) found that activity in the ventral striatum, which is typically associated with wanting rather than liking, was larger on self-initiated joint-attention trials, which were also rated more pleasant than non-shared attention trials. Perhaps liking effect participants are more motivated to share attention and therefore are more likely to follow others' gaze and be susceptible to gaze-cueing effects. Subsequently, the pleasant feelings on congruent trials and lack of pleasant feelings on incongruent trials may influence liking ratings. However, the interaction between validity and liking effect in reaction times was not observed in experiment 5 and may need investigating further before more solid conclusions can be formed. Note also that any differences between this and other experiments could also be due to the use of a new stimulus set or faces with no pupil manipulation in the case of ERP results. However, there were still gaze-cueing effects and similar initial liking ratings making this unlikely.

In conclusion, experiment 6 has investigated the reliability of liking effects from identity-contingent gaze-cues and whether the ERPs elicited in the context of liking are similar to those elicited in the context of trust. While the explicit liking ratings compound with the results of all previous experiments showing that liking ratings relate less, or in this case, not at all, to evaluations of faces based on identity-contingent gaze-cues, participants implicit emotional responses to the gaze-cues were generally in the same direction as that seen with trust in experiment 1 in that they were larger for the more negative incongruent faces. However, there were many differences in the details of the effects such as the strength, polarity, spatial distribution and location of peaks. This means that there probably are some ways in which the neural representation and processes by which emotion interacts with and recruits cognitive processes differ depending on whether participants are primed with liking compared to trustworthiness in the initial ratings. The only way of understanding this may be to use fMRI to image the detailed interaction between different functional networks involved in emotion, attention, perception and memory. It may also be useful to measure both trustworthiness and liking ratings in the same task to make sure that both types of ratings have the potential to influence how gaze-cues are processed and to investigate whether trust relates to liking ratings.

# Chapter 8

## General Discussion

In this thesis, a series of six experiments has attempted to understand how identity-contingent gaze-cues elicit changes in trustworthiness judgements of faces. Trustworthiness entails accepting risk upon another person's behalf (Rousseau et al., 1998). Trusting another person can mean the difference between positive, rewarding outcomes and negative, punishing outcomes. Indeed, previous research has shown that neural and psychophysiological correlates of emotion are recruited when judging (Adolphs et al., 1998; Engell et al., 2007; Marzi et al., 2014; Winston et al., 2002; Yang et al., 2011) and learning (King-Casas et al., 2005; Hofman et al., 2012) the trustworthiness of other individuals and observing gaze-cues (Gordon et al., 2013; Schilbach et al., 2010). Therefore we hypothesised that emotion and somatic markers mediate the learning of trust from gaze-cues. Experiments 1 and 2 investigated this by measuring psychophysiological correlates of emotion from the brain and facial muscles during identity-contingent gaze-cueing. Experiment 3 took a slightly different approach by looking at whether priming the emotions associated with social exclusion/ inclusion modulates the trust effect. Experiments 4, 5 and 6 were follow up experiments which aimed to investigate whether the learning of trust from gaze-cues conforms to prediction-error learning and if the effects observed previously are specific to faces, gaze and trust judgements or generalize to arrows and liking.

Overall, we argue the findings conform to our hypotheses. In experiment 1, the late positive potential (LPP) ERP, which is related to emotion processing, was sustained throughout the trial, after emerging in the final 500 ms of trial period 2, and increased to incongruent faces as blocks progressed. The interaction of validity with block suggested that emotion is not only elicited by gaze-cues but also implicated in learning evaluations, which are subsequently retrieved when observing the face with neutral gaze. Although some research has shown that participants can make accurate judgements of trustworthiness with as little as 100 ms exposure to a face (Willis & Todorov, 2006), it appears that emotion based evaluations take at least 1000 ms to emerge after face onset. In experiment 2, evidence for the role of somatic markers of emotion was demonstrated in corrugator facial EMG activity and appeared to relate to trust judgements of faces in the predicted directions. Corrugator activity in trial period 4 was greater for incongruent compared to congruent gaze-cues only in those participants who showed a trust effect in the ratings. Somewhat unexpectedly, although not completely inconsistent with our hypotheses, participants who did not show a trust effect also showed the reverse pattern in trial periods 2 and 3 where corrugator activity was greater to congruent compared to incongruent faces. In experiment 3, we found no significant effects of the priming of the social emotions associated with social inclusion or exclusion on EMG activity and learning of trust from congruent and incongruent identity-contingent gaze-cues, although the latter appeared to be in the anticipated direction of a larger trust effect for participants primed with exclusion. This may be because rejection and trustworthiness are distinct social emotions or because the source of the emotions were highly salient and therefore was less likely to generalise from one to the other. The fact that social exclusion/ inclusion did not produce an effect in both the trust ratings or EMG activity further illustrates the relation between the two measures. However, unlike experiment 2, interactions between trust effect and validity in EMG activity were not observed despite the same main effects of validity in trial period 4 in the corrugator and zygomaticus. This could be related to influential companions, the phenomenon whereby effects change depending on the introduction of other variables into the experimental design (Poulton, 1982).

It is unclear exactly what it is about the trust effect participants in experiment 2 that differentiates them from the no trust effect participants in terms of EMG reactivity to gaze-cues and their influence on ratings. We did not detect any differences in anxiety between trust effect and no trust effect participants or any effects on gaze-cueing, trust and EMG activity. We also did not find any effects of autism traits using the AQ and empathy with the EQ in experiment 3. However, it remains to be demonstrated whether depressive tendencies interact with the trust effect and EMG. Indeed, a previous unpublished experiment has shown that participants with a history of depression do not learn to mistrust incongruent faces to a greater extent than congruent faces in comparison with matched participants without a history of depression. This is despite equivalent cueing effects (Rogers, Bayliss, Wakeley, Cowen & Tipper, unpublished manuscript). However, it was not simply the absence of an emotional response that drived EMG some effects. In experiment 2, we found that no trust effect participants showed a reverse trust effect where they rated congruent faces less trustworthy than incongruent faces and showed larger corrugator EMG activity to the former compared to the latter in trial periods 2 and 3. Interestingly, these reverse effects are consistent with Aguado et al. (2013), which found that participants who did not show EMG mimicry of identity-contingent angry and happy faces also showed patterns of EMG consistent with reverse mimicry when the faces were presented in a neutral state before facial expression after several blocks of learning. The reverse effects may well be due to personality attributes that are not easily captured by questionnaires. As mentioned previously, one possible explanation could be that no trust effect participants are less sensitive to cues of deception because they are deceptive themselves and accept this as a moral norm. However, trustworthy congruent cues may be experienced as emotionally negative as it is uncomfortable to deceive an individual that has behaved trustworthy. It may be useful for future research to examine the large variety of different personality measures that could relate to the learning of trust from gaze such as schizotypy, narcissism and the big five personality traits.

We also did not find any differences between trust effect and no trust effect participants in experiment 1 measuring EEG. One suggestion for future investigation would be to divide participants based on resting state frontal alpha EEG asymmetry. Greater left than right asymmetry has been shown to be related to trait positive affect and increased positive affect to positive stimuli whereas the inverse is associated with trait negative affect and larger negative affect to negative stimuli (Tomarken et al., 1992; Wheeler et al., 1993). It may also be worth taking a similar approach with EMG. Instead of dividing participants based on trust effect and examining differences between groups, it may be more useful to divide participants based on whether they show differences between conditions based on EMG activity and then examine differences in ratings between groups.

Experiments 1-3 showed that the learning of trust from congruent and incongruent gaze-cues was modulated by face gender. The learning of mistrust from incongruent gaze-cues was larger for female faces and the learning of trust from congruent gaze-cues was larger for male faces. This was evidenced by an interaction of time with gender. In experiment 1, the interaction was driven by the final ratings whereas in experiments 2 and 3 the effect was driven by the initial ratings. This suggested that the learning of trust is greater when there is a mismatch between how trustworthy a person is expected to behave and how trustworthy the person actually behaves. Such a pattern is consistent with prediction-error learning rules such as that proposed by Rescorla and Wagner (1972). Therefore, in a stronger test of this hypothesis, experiment 4 examined if the effects of initial expectancies on learning of trust from gaze-cues was greater when faces are much more highly polarised in terms of initial trust ratings. Surprisingly, although the effects were in the anticipated directions, the learning of trust from gaze-cues was not significantly modulated by initial trustworthiness. As in experiment 3, if the sources of participants trust judgements from facial appearance are highly salient there may be less generalisation to evaluations based on gaze or such judgements in the initial ratings may interfere with fine grained learning.

Alternatively, the learning of trust may rely upon evaluative conditioning learning mechanisms such as referential (Baeyens et al., 1992), holistic (Martin & Levey, 1978; 1994) or implicit misattribution learning (Jones et al., 2009), which do not rely upon prediction-errors but could account for the interactions with gender. According to the referential account, learning occurs when stimuli occur in close proximity in time and space. The neural mechanisms underpinning this type of learning could conform to Hebbian (1949) learning according to which neurons that fire together wire together. In this case neurons that represent face identity become associated with neurons that represent emotional states. Subsequently, activation of one of these representations spreads to the representation of the other. In this way, learning does not depend on the prediction error. The fact that we detected retrieval effects in the LPP in experiment 1 would seem to support this. The holistic account makes similar predictions although the representation is slightly different. Instead of being independent but linked, face identity and emotion become apart of the same representation and therefore the emotional states generalise to face identity. However, we did not detect any effects of validity on the face-recognition related N250 ERP component, which might rule out this explanation. The implicit misattribution learning account states that, because conditioned and unconditioned stimuli occur in close spatial and temporal proximity, the emotional response to the latter generalises to the former. This is highly plausible as the eyes are the most important features for face perception and recognition (Sormaz, Andrews & Young, 2013), are the sources of the evoked emotion and occur in the context of the face.

Experiment 5 showed that the learning of evaluations from gaze-cues is not due to domain general attentional cueing as the effect did not generalise to arrows which cue attention in a similar way to gaze and which were found to be equally as distinctive and memorable in a preliminary visual search and memory task. Although both types of cue elicited activity in the corrugator and zygomaticus in trial period 4, this activity was larger and more sustained as evidenced by the larger response to incongruent compared to congruent faces but not arrows in trial period 5. This suggests, that while there may be some emotional reaction to attention misdirection by arrows, this does not appear to be as strong or sustained as that elicited by gaze and does not appear to elicit changes in liking. As the effect for faces was sustained from trial period 4 to 5, the learning of liking may be related to evaluative iterations where reprocessing of initial automatic implicit evaluations leads to more reflective explicit judgements that integrate contextual information, such as face identity, into the evaluation (Cunningham & Zelazo, 2007). However, the liking effect was weaker than that observed in previous experiments with trustworthiness and was driven primarily by congruent faces. This special relationship between trustworthiness, faces and gaze-cues was confirmed in experiment 6 which failed to find an effect of gaze-cues on the learning of liking despite significant effects of validity throughout the whole trial and interactions with block in trial period 2 in the LPP. These ERP effects differed in latency and scalp distribution to experiment 1 suggesting a potential influence of the initial liking ratings, although the anterior laterality of the peaks suggested that the valence of the response was the same. Such findings suggest that implicit emotional reactions to gaze-cues are independent of explicit evaluations but can inform them and be informed by them.

In combination with experiment 3, the findings of experiment 5 and 6 suggest that trust may be a distinct social emotion, being elicited by specific representations of others thoughts, feelings or actions, in this case identity-contingent gaze-cues (Hareli & Parkinson, 2008). It will be interesting for future research to determine whether the effects are due to the physical properties of the gaze-cueing procedure or if it is driven by the mental state attribution of intentions and motivations to deceive or share attention. Even low-level gaze-cueing is modulated by mental state attribution (Teufel et al., 2010). One potential way in which this can be studied is by using a paradigm in which the head shifts unpredictably at the time of the gaze-cue which can make the gaze shift appear as if it was due to a head movement in the opposite direction instead of a voluntary eye-movement whilst maintaining cueing effects (Bayliss, di Pellegrino & Tipper, 2005; Rogers et al., 2014).

The somatic marker hypothesis was key to our predictions (Damasio, 1996). This theory states that feedback from embodied emotion states is used to infer emotion and guide decision making and cognition. Our EMG findings from experiments 2, 3 and 5 suggest a role for somatic markers in the learning of trustworthiness from gaze-cues. In these experiments larger activity was observed in response to incongruent gaze-cues in the corrugator and zygomaticus muscles both of which have been shown to increase in response to emotionally arousing negatively valenced affective pictures (Greenwald et al., 1989; Lang et al., 1993; Larsen et al., 2003). However, only in experiment 2 were there interactions with individual differences in trust evaluations. As mentioned previously, this may be because of influential companion effects in experiment 3 and the fact that liking ratings relate less strongly to face evaluation than trust in experiment 5 (Oosterhof & Todorov, 2008). The fact that the basic effect was observed in both experiments 2 and 3, which measured trust, and also in experiment 5, which measured liking, suggests that it is an immediate and implicit response to gaze-congruency. Although this same reaction also appears to be evident in response to arrow congruency this does not appear to elicit changes in cognition that lead to learning of liking. As EEG may better demonstrate such changes it would be valuable for future research to measure EEG during the arrow-cueing experiment to confirm this.

We chose EEG and EMG as implicit measures of emotion. The corrugator muscle is more responsive to stimuli of negative valence whereas the LPP and zygomaticus muscle can respond to stimuli of both positive and negative valance (Cuthbert et al., 2000; Greenwald et al., 1989; Lang et al., 1993; Larsen et al., 2003; Schupp et al., 2000). Throughout all of our experiments we detected changes in EMG and EEG that indicated the negative, untrustworthy, incongruent faces were responded to more strongly than the positive, trustworthy, congruent faces. As all of the important effects were found in the corrugator and the effects in the zygomaticus were not found without effects in the corrugator this suggests the changes in EMG were due to negative emotions. In ERP experiment 1, the validity and block interaction peaked in the right posterior hemisphere, where LPP effects are typically measured, and increased to incongruent faces across blocks. In experiment 6, the LPP was larger to incongruent on right frontal electrodes where larger activity is typically observed for negative emotions (Harmon-jones et al., 2010; Tomarken et al., 1992; Wheeler et al., 1993). Therefore, overall, the findings are consistent with the notion that participants are more sensitive or expressive towards the incongruent faces. This is consistent with studies showing negativity biases in the LPP (Hilgard et al., 2014; Ito et al., 1998; Wood & Kisley, 2006), where the LPP is larger to negative compared to positive stimuli which are equivalent in terms of the absolute value of the difference in valence and arousal ratings to neutral stimuli. The LPP is also greater for faces judged untrustworthy compared to trustworthy (Marzi et al., 2014; Yang et al., 2011). The negativity bias in the LPP and EMG response to incongruent faces and its relation to trustworthiness may be similar to loss aversion, where participants are less likely to make decisions in which there is a chance that a loss may occur, or when the choice is framed in terms of loss, compared to when there is an equal possibility that a gain of equivalent magnitude and probability could occur (Tversky & Kahneman, 1984). In this case, the loss is in terms of being misled to the location of the object and the omission of joint-attention. As the corrugator does not increase to positive stimuli, and even shows decreases to such stimuli (Larsen & Norris, 2009), it makes sense that significant effects were found more frequently in this muscle. The differences between incongruent and congruent conditions in trust ratings and EMG activity may have been facilitated by contrast between the two. Such affective contrast has previously been demonstrated in emotion-related EMG activity (Larsen & Norris, 2009) and trustworthiness judgements (Wincenciak et al., 2013), highlighting a beneficial role for influential companions (Poulton, 1982). The findings of experiment 1 are consistent with Schilbach et al. (2010), which showed larger pupil size during non-joint attention, which, like the LPP, is related to emotional arousal (Bradley et al., 2008).

Evaluative ratings were measured both before and after cueing in all experiments. The acquisition of ratings before cueing provided an initial baseline to compare with after cueing in order to control for any initial variability across conditions and identify the direction of any effects. Despite these benefits it could also potentially bias the way in which the stimuli are subsequently processed in the cueing and/ or viewing phases. As the EMG and EEG data were consistent across experiments in suggesting that the incongruent faces generally elicit more negative emotions than congruent faces, basic implicit emotional responses appear to be unaffected by the initial ratings, although scalp distributions and timing may vary. However, the effects of gaze-cues on liking ratings were weaker or non-existent in experiments 5 and 6 compared to the strong effects in the trust ratings seen in experiments 1 to 4. In the one experiment (5), which did show effects of gaze-cues on liking, the effect was driven primarily by congruent faces. This suggests that liking judgements may engender a focus on the extent of the positive attributes of evaluated stimuli without inclusion of the negative attributes. It remains to be demonstrated whether the differences between liking and trust ratings is because of semantic differences or also because of the effects of priming these concepts in the initial ratings on subsequent processing during cueing.

Another aspect of our design that may have contributed to the observed effects is the initial eye contact that is established in trial period 2. Eye contact compared to aversion elicits larger SCRs, particularly for smiling faces, and individual differences in frontal alpha-asymmetry is correlated with alpha-asymmetry differences in response to gaze and aversion such that individuals with greater left frontal activity show larger left frontal activity in response to eye contact compared to aversion (Pönkänen & Hietanen, 2012). As mentioned previously, SCRs are associated with emotional arousal (Cuthbert et al., 2000) and left frontal EEG activity is associated with positive affect (Harmon-Jones et al., 2010). It would be useful to examine the effects of initial eye contact on the learning of trust from gaze-cues. One might predict that the effects of incongruent gaze-cues on negative emotion and trust are larger when an initial positive emotion has been established through eye contact.

We did not assess participant's conscious awareness of the contingencies between face identity and gaze-cues. However, this issue has been addressed by similar studies (Bayliss et al., 2009; Rogers et al., 2014) where pairs of matched faces were presented to participants after gaze-cueing. One of the faces had always gazed congruently whereas the other had always gazed incongruently. When participants judged which of the two faces was more likely to look towards the target, they chose incongruent faces equally as frequently as congruent faces. Throughout all experiments, we also found the gaze-cueing effect was still evident in block 5 after repeated exposures to the identity-contingent gaze-cues. This suggests that participants were not using the identity and gaze direction of the face to anticipate target location, which would be expected if participants had conscious knowledge of the contingencies between identity and gaze-cues. Also, participants are typically poor at remembering specific features of faces (Tanaka & Farah, 1993) compared to the complex conjunction of features that form holistic representations necessary for recognition that is indexed by the N250 effects. However, we do not make strong claims concerning awareness of gaze contingencies. There were changes to the procedure, such as measurement of initial trust ratings, and we did not explicitly investigate the awareness issue.

One limitation of the experiments investigating effects of identity-contingent gaze-cues on trust is related to the way in which trust is measured. Trustworthiness in relation to gaze-cues can be construed in one of two ways. One way it is most commonly understood, and construed in this thesis, is in terms of the general willingness to accept risk on another persons behalf regardless of whether that is loaning another person a large amount of money or depending on someone to belay your rope whilst climbing a large rock face. On the other hand, in the current series of experiments trust can also be construed more specifically in terms of whether one can rely on the gaze-cues to direct the location of the target. The available experimental evidence would seem to suggest that the former construal of trust is more likely learned from gaze-cues. Learning of trust from identity-contingent gaze-cues generalises from trust judgements to trust behaviour in other domains, including in economic exchanges (Rogers et al., 2014). For future research we suggest the use of well established and validated specific interpersonal trust measures that ask participants to answer a range of questions about their willingness to trust a specific other person in a variety of different circumstances and has subscales such as reliability and emotional trust which are distinct from liking (Johnson-George & Swap, 1982). In addition, instead of dividing participants in terms of whether they show trust effects, it may also be valuable to divide participants based on scores on the interpersonal trust scale devised and validated by Rotter (1967; Robinson, Shaver & Wrightsman, 1991). This questionnaire assesses an individual's general propensity to trust or mistrust other people. As such, one may predict that mistrusting people are more sensitive to incongruent gaze-cues whereas trusting people are more sensitive to congruent gaze-cues (Gurtman & Lion, 1982). Effects such as these could suggest potential utility of dividing participants trust effects in more complex ways such as depending on whether participant's ratings are generally negative, positive or opposite in polarity depending on the gaze-cue rather than simply on the sign of the difference between congruent and incongruent. More consistent trust effects may be found by examining the relation between individual differences in trust behaviours (as in Rogers et al.) and EMG activity instead of trust ratings.

In all of the experiments conducted, the participant's task was to classify objects as to whether they belong in a kitchen or a garage. While this task is well suited to examining gaze-cueing effects it is not particularly reflective of real life behaviour, as the kitchen and garage objects have no importance other than task relevance. Incongruent gaze-cues can be used to mislead others attention, but this is presumably more likely to be used when the stimulus has high value to the gazer or the individual being deceived. The task could be changed to ensure that the detection of the target is more important for the participant. There are many different ways in which this can be achieved. For example, if the target stimuli were money and its acquisition was dependent upon, or weighted by, reaction time, there may be stronger trust effects as the cost in reaction times due to incongruent cues would be greater. This might be particularly the case if the money lost by the participant is believed to be gained by the gaze-cueing face. There is some evidence that gaze-cueing is modulated by the emotional relevance of the target stimulus. Kuhn and Tipples (2011) found that participants tended to follow the gaze of fearful faces more than happy faces, as measured by eye tracking, when participants searched for a target that was threatening as opposed to pleasant. Bayliss, Schuch and Tipper (2010) also found that the gaze-cueing effect was larger for faces that expressed happiness as opposed to disgust for participants who had to detect targets that were pleasant IAPS images. No interactions were found for negative IAPS images. Therefore one might predict similar interactions between trustworthiness judgements of faces and identity-contingent gaze-cues that lead to monetary gains or losses.

It may also be useful for future research to measure trust ratings of faces continuously throughout the experiment such as in between blocks in order to ascertain exactly when differences in trust ratings emerge and if this is similar to the linear LPP effect observed in experiment 1. This could help identify whether trust evaluations are learned rapidly and are categorical or gradual. That is, the magnitude of the trust rating could depend on the strength of the emotional response, which accrues over time as the face is repeatedly observed. Alternatively, participants could learn trust on the first trial and simply categorise the face as trustworthy or untrustworthy. The findings of experiment 1 suggest that the emotional response to incongruent faces gradually increases across blocks. Whether the same applies to explicit trust judgements remains to be demonstrated. Indeed, one possible reason why there was no interaction between face trustworthiness and validity in experiment 4 could be because the larger changes in trust ratings for high trust incongruent faces and low trust congruent faces occurs rapidly within a small number of trials, diminishing thereafter. In all experiments we measured trust once per face before and after cueing. Future research could take ratings several more times in order to obtain a more reliable and accurate rating and to examine EMG and EEG activity during the rating trials.

Throughout all experiments, the validity of the gaze-cues was identity-contingent. We compared congruent faces with incongruent faces because one could expect that congruent faces are judged positively and incongruent faces are judged negatively in terms of trust and indeed this pattern was observed across experiments. Therefore comparing these two conditions, which differ significantly in terms of trust, maximised our chances of detecting a contrast in the EMG and EEG. However for future research it may be valuable to use a control condition in which a third face-validity category is used where particular faces are presented that either do not gaze-cue at all or gaze-cue congruently and incongruently an equal number of times. The use of a non-gazing face control condition may be more useful for detecting effects of emotion on EMG and EEG activity. As the LPP and the zygomaticus muscle tend to respond to both positive and negatively valenced emotional stimuli, comparison with a neutral condition may be more likely to detect effects and show effects that are specific to incongruent or congruent faces and not just a product of comparing the difference between the two. Indeed, neuroimaging studies have shown that the observed patterns of activation in the amygdala in response to emotional stimuli depend on the type of baseline control comparison stimulus used (Mattavelli et al., 2014). The use of a control condition in which faces gaze congruently and incongruently an equal number of times would be useful for examining reversals in face congruency between successive trials which may provide another line of inquiry into whether the learning of trust depends on prediction-errors. It may also be valuable to vary the ratio of the number of trials on which faces gaze congruently or incongruently to examine the effects of volatility and if participants are biased to base their judgements on incongruent cues when both types of cues are emitted by face.

In all of our experiments it is obvious to participants that the gaze behaviour of the faces are controlled by the computer. It would therefore be valuable to examine the effects of identity-contingent gaze-cues on trust judgements in a more realistic and natural context. One way in which this has been studied previously is by using gaze-contingent eye tracking and misleading participants into believing that the face observed on the screen is being controlled by another participant's eye-movements (Pfeiffer, Vogeley & Schilbach, 2013; Schilbach et al., 2010; Wilms et al., 2010). The believability that the gaze of the face is being controlled by a real person could also be increased by using videos of gaze-shifts that participants believe are being transmitted in real-time via cameras. The ideal approach would be to use real participants interacting in a face-to-face context, an approach that has recently been successfully used to demonstrate basic gaze-cueing effects in real life interactions (Lachat et al., 2012a). However, due to the need for multiple faces in the incongruent and congruent conditions, this may not be feasible. It may also be worth examining whether a reaction time cost is necessary for the trust effect as opposed to simply the perception of incongruent gaze as gaze-cueing effects did not differ between trust effect and no trust effect groups.

Although, we did not perform source localisation on effects in the cueing task, one can make predictions of the neural systems that may be mediating the effects based on previous research. In source localisation and combined EEG-fMRI studies, the LPP has been related to activity in brain regions involved in visual/ attentional processing such as lateral occipital, parietal and inferotemporal cortices and emotion regions such as the OFC, insula, anterior cingulate cortex, ventral striatum and amygdala (Harris et al., 2011; Liu et al., 2012; Moratti et al., 2011; Sabatinelli et al., 2007; 2013). The amygdala is also correlated with corrugator activity in response to negative affective pictures (Heller et al., 2011; 2014) and plays a role in enhancing visual perception and attention in response to emotion stimuli (Vuilleumier & Driver, 2007). Therefore the amygdala may be involved in evaluating faces/ gaze-cues and coordinating the emotional response by triggering facial muscle activity and the attentional changes related to the LPP. Indeed there is a wealth of evidence from a range of methods showing amygdala responses to eyes, eye contact and gaze-cues (Huijgen et al., 2015; Mosher et al., 2014; Sato et al., 2011). This could also explain the consistent negativity biases as amygdala activity is greater to social cues that are more negative and threatening (Amaral, 2003; Mattavelli et al., 2014). The changes in attention may be reflected in increased activity in visual areas important for processing faces (such as OFA, STS and FFA) and gaze (aSTS, pSTS, IPS). However, the ventral striatum has been more implicated in learning trust than the amygdala (e.g. King-Casas et al., 2005). The striatum may also mediate initial motivation to learn face trustworthiness (Pfeiffer et al., 2014) and engage in joint attention (Gordon et al., 2013; Schilbach et al., 2010). The importance of initial motivation may be evident in experiment 1 where there was generally faster reaction times for trust effect participants and in experiment 5 where there was increased corrugator activity for arrows, which did not subsequently show evidence of learning of liking.

There were several potential reasons why identity-contingent gaze-cues elicited emotion and changes in trust ratings. Firstly, people are motivated to interact with other people and to share attention and therefore the joint-attention inherent to congruent cues may be experienced as more pleasant than the non-joint attention of incongruent cues (Tomasello et al., 2005). Secondly, as congruent gaze cues lead to faster reaction times than the incongruent cues, this may elicit increased visuomotor fluency for congruent compared to incongruent faces, which has been shown also to elicit more positive affect including increased liking and zygomatic EMG activity (Cannon et al., 2010; Regenberg et al., 2012; Winkielman & Cacioppo, 2001; Winkielman et al., 2006). Thirdly, the effects could be due to violation of expectancy by incongruent faces. Finally, the effects could be related specifically to the deceptiveness of the gaze-cues. These explanations may not be mutually exclusive, but the strongest contribution may come from the relationship between gaze-cues and deception/ trust. This is because gaze-cues had a stronger effect on ratings of trust than liking and because the emotional reaction in experiment 2 appeared to relate to trust ratings.

However, there may be a role for fluency in the learning of liking. Liking is the type of evaluative rating typically measured in studies of fluency. In experiment 5 the EMG response in trial period 5 was larger to incongruent compared to congruent faces but not arrows and only the former showed a change in liking. Also, the cueing effect in experiment 6 was larger for liking effect participants compared to non-liking effect participants. It is difficult to know the influence of the pleasantness of joint-attention on trust and liking judgements. One might predict that if this were the case, participants might prolong the duration of the joint-attention episode and therefore show faster responses to incongruent gaze-cues due to the tendency to look longer at preferred stimuli (Schotter, Berry, McKenzie & Rayner, 2010). Although the participants' task was to detect the object as quickly as possible, this is also not supported by eye-tracking studies, which have shown that congruent faces tend to be gazed at for a shorter duration than incongruent faces in early trial periods after several blocks of learning (Bayliss et al., 2013). Violation of expectancy does not appear to explain the effects as incongruent arrow cues violate expectancy to the same degree as incongruent faces/ gaze cues but did not elicit changes in liking. Also, EMG activity to faces was more discriminative of validity in trial period 5. However, the effect of validity for arrows in trial period 4 could be an immediate response to expectancy violation. This would also apply to faces and is a better explanation of this effect than in terms of joint-attention, as arrows do not elicit joint-attention states that are specific to faces and gaze. In the case of faces, the violations may be in terms of expectancies about trustworthiness. The importance of expectancies was highlighted by the face gender effects in experiments 2 and 3. It was also suggested by the larger LPP to congruent faces in early blocks of experiment 1.

We also must point out that there are alternative explanations to the notion that the emotional reactions measured with EMG and EEG are due specifically to the deceptiveness of the gaze-cues. Many studies have shown that objects that are consistently gazed towards are liked more than objects that are consistently gazed away from in the gaze-cueing task (Bayliss et al., 2006; 2013; Capozzi et al., 2014; Manera et al., 2014; Treinen et al., 2012; Ulloa et al., 2015). The EMG and EEG responses could possibly reflect the embodied simulation/ representation of another person's dislike of objects and the learning of preferences or others' preferences. This could only partially explain the findings as strong trust effects were observed in experiments 1-4 and EMG activity was related to trust ratings in experiment 2, although a relationship between trust and liking of objects from gaze-cues has been demonstrated. The increased liking of objects gazed towards compared to away is greater for faces judged trustworthy compared to untrustworthy (Treinen et al., 2012). There are also other interpretations. As the incongruent faces in our paradigm did not look at the objects and in fact averted gaze before the target was presented, in a way the incongruent gaze-cueing face is also being misled in that they are perceived as not looking at the object. Therefore the emotional responses could potentially reflect the interpretation of incongruent faces as being ignorant or uninformed of the location of an important object.

In conclusion, in this series of six experiments, we have provided novel EEG and EMG evidence for a mediating role of emotion in the learning of trust from identity-contingent gaze-cueing. This appears to be due to a specific relationship between faces, gaze, emotion and trust. As such, our findings provide further support for the influence of emotion and embodied states on social cognition and suggest several potentially valuable avenues for future investigation. Distraction and diversion are the most important tools used by street thieves who steal from unsuspecting others in the crowded social environments that characterize cities across the world. The use of eye-gaze to mislead attention away from such crimes is vividly illustrated in detailed paintings, such as *The fortune teller, The cardsharps* and *The cheat with the ace of diamonds,* created by Michelangelo Merisi da Caravaggio and Georges de La Tour around 400 years ago. Perhaps, all those years ago, these artists were attempting to elicit in observers the kinds of emotions we have identified in response to similar cues in our paradigm as being important for learning trustworthiness.

## References

Adolphs, R., Tranel, D., & Damasio, A.R. (1998). The human amygdala in social judgment. *Nature, 393*(6684), 470-474. doi: 10.1038/30982

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*(6507), 669-672. doi: 10.1038/372669a0

Aggleton, J.P., & Young, A.W. (2000). The enigma of the amygdala: On its contribution to human emotion. In R.D. Lane & L. Nadel (Eds.), *Cognitive Neuroscience of Emotion* (pp. 242-276). New York: Oxford University Press.

Aguado, L., Román, F.J., Fernández-Cahill, M., Diéguez-Risco, T., & Romero-Ferreiro, V. (2011). Learning about faces: effects of trustworthiness on affective evaluation. *Spanish Journal of Psychology, 14*(2), 523-534. doi: 10.5209/rev_SJOP.2011.v.14.n2.1

Aguado, L., Román, F.J., Rodríguez, S., Diéguez-Risco, T., Romero-Ferreiro, V., & Fernández-Cahill, M. (2013). Learning of facial responses to faces associated with positive or negative emotional expressions. *Spanish Journal of Psychology, 16*(e24), 1-10. doi: 10.1017/sjp.2013.31

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences, 4*(7), 267-278. doi: 10.1016/S1364-6613(00)01501-1

Allison, T., Puce, A., Spencer, D.D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. 1:Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9*(5), 415-430. doi: 10.1093/cercor/9.5.415

Amaral, D.G. (2003). The amygdala, social behaviour and danger detection. *Annals of the New york Academy of Sciences, 1000,* 337-347. doi: 10.1196/annals.1280.015

Amemiya, S., & Ohtomo, K. (2012). Effect of the observed pupil size on the amygdala of the beholders. *Social, Cognitive and Affective Neuroscience, 7*(3), 332-341. doi: 10.1093/scan/nsr013

Anderson, N.C., Risko, E.F., & Kingstone, A. (2011). Exploiting human sensitivity to gaze for tracking the eyes. *Behavioral Research Methods, 43*(3), 843-52. doi: 10.3758/s13428-011-0078-8

Andrews, T.J., Davies-Thompson, J., Kingstone, A., & Young, A.W. (2010). Internal and external features of the face are represented holistically in face-selective regions of the visual cortex. *Journal of Neuroscience, 30*(9), 3544-3452. doi: 10.1523/JNEUROSCI.4863-09.2010

Anzellotti, S., & Caramazza, A. (2014). The neural mechanisms for the recognition of face identity in humans. *Frontiers in Psychology, 5*(672), 1-6. doi: 10.3389/fpsyg.2014.00672

Ariely, D., & Norton, M.I. (2008). How actions create- not just reveal- preferences. *Trends in Cognitive Sciences, 12*(1), 13-16. doi: 10.1016/j.tics.2007.10.008

Armel, K.C., Pulido, C., Wixted, J.T., & Chiba, A.A. (2009). The smart gut: Tracking affective associative learning with measures of “liking’, facial electromyography, and preferential looking. *Learning and Motivation, 40,* 74-93. doi: 10.1016/j.lmot.2008.06.003

Bach, P., & Tipper, S.P. (2007). Implicit action encoding influences personal-trait judgments. *Cognition, 102*(2), 151-178. doi: 10.1016/j.cognition.2005.11.003

Baeyens, F., Eelen, P., Crombez, G., & Van den Bergh, O. (1992). Human evaluative conditioning: acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy, 30*(2), 133-142. doi: 10.1016/0005-7967(92)90136-5

Ballantine, H.T., Cassidy, W.L., Flanagan, N.B., & Marino, R. (1967). Stereotaxic anterior cingulotomy for neuropsychiatric illness and intractable pain. *Journal of Neurosurgery, 26*(5), 488-495. doi: 10.3171/jns.1967.26.5.0488

Baron, S.G., Gobbini, M.I., Engell, A.D., & Todorov, A. (2011). Amygdala and dorsomedial prefrontal cortex responses to appearance-based and behavior-based person impressions. *Social, Cognitive & Affective Neuroscience, 6*(5), 572-581. doi: 10.1093/scan/nsq086

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind.* Cambridge: MIT Press.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*(2), 163-175. doi: 10.1023/B:JADD.0000022607.19833.00

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

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-609.

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs: Oscillatory Neuronal Dynamics. In S.J. Luck & E.S. Kappenman (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 31-49). New York: Oxford University Press. doi: 10.1093/oxfordhb/9780195374148.001.0001

Baxter, M.G., & Murray, E.A. (2002). The amygdala and reward. *Nature Reviews Neuroscience, 3*(7), 563-573. doi: 10.1038/nrn875

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*(4), 634-644. doi: 10.1037/a0020559

Bayliss, A.P., di Pellegrino, G., & Tipper, S.P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *The Quarterly Journal of Experimental Psychology, 58A*(4), 631-650. doi: 10.1080/02724980443000124

Bayliss, A.P., Griffiths, D., & Tipper, S.P. (2009). Predictive gaze cues affect face evaluations: The effect of facial emotion. *European Journal of Cognitive Psychology, 21*(7), 1072-1084. doi: 10.1080/09541440802553490

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), 76-92. doi: 10.1037/a0029286

Bayliss, A.P., Paul, M.A., Cannon, P.R. & Tipper, S.P. (2006). Gaze cueing and affective judgments: I like what you look at. *Psychonomic Bulletin & Review*, 13, 1061-1666. doi: 10.3758/BF03213926

Bayliss, A.P., Schuch, S., & Tipper, S.P. (2010). Gaze cueing elicited by emotional faces is influenced by affective context. *Visual Cognition, 18*(8), 1214-1232. doi: 10.1080/13506285.2010.484657

Bayliss, A.P., & Tipper, S.P. (2005). Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *British Journal of Psychology, 96*(1), 95-114. doi: 10.1348/000712604X15626

Bayliss, A.P., & Tipper, S.P. (2006). Predictive Gaze Cues and Personality Judgements: Should Eye Trust You? *Psychological Science, 17*(6), 514-520. doi: 10.1111/j.1467-9280.2006.01737.x

Bechara, A., Damasio, H., Tranel, D. & Damasio, A.R. (1997). Deciding Advantageously Before Knowing the Advantageous Strategy. *Science, 275,* 1293-1295. doi: 10.1126/science.275.5304.1293

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A.R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science, 269*(5227), 1115-1118. doi: 10.1126/science.7652558

Beckers, T., De Vicq, P., & Baeyens, F. (2009). Evaluative Conditioning is Insensitive to Blocking. *Psychologica Belgica, 49(1),* 41-57. doi: 10.1080/17470218.2014.939667

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

Belova, M.A., Paton, J.J., & Salzman, C.D. (2008). Moment-to-moment tracking of state value in the amygdala. *Journal of Neuroscience, 28*(40), 10023-10030. doi: 10.1523/JNEUROSCI.1400-08.2008

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. doi: 10.1162/jocn.1996.8.6.551

Berridge, K.C., & Robinson, T.E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews, 28*(3), 309-369. doi: 10.1016/S0165-0173(98)00019-8

Bindemann, M., Burton, A.M., Leuthold, H., & Schweinberger, S.R. (2008). Brain potential correlates of face recognition: geometric distortions and the N250r brain response to stimulus repetitions. *Psychophysiology, 45*(4), 535-544. doi: 10.1111/j.1469-8986.2008.00663.x

Bisley, J.W., & Goldberg, M.E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience, 33,* 1-21. doi: 10.1146/annurev-neuro-060909-152823

Black, F.W. (1975). Unilateral brain lesions and MMPI performance: a preliminary study. *Perceptual and Motor Skills, 40,* 87-93. doi: 10.2466/pms.1975.40.1.87

Blackhart, G.C., Nelson, B.C., Knowles, M.L., & Baumeister, R.F. (2009). Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: a meta-analytic review of 192 studies on social exclusion. *Personality and Social Psychology Review, 13*(4), 269-309. doi: 10.1177/1088868309346065

Boutsen, L., Humphreys, G.W., Praamstra, P., & Warbrick, T. (2006). Comparing neural correlates of configural processing in faces and objects: an ERP study of the Thatcher illusion. *Neuroimage, 32*(1), 352-367. doi: 10.1016/j.neuroimage.2006.03.023

Bradley, M.M., Greenwald, M.K., Petry, M.C., & Lang, P.J. (1992). Remembering pictures: pleasure and arousal in memory. *Journal of Experimental Psychology: Learning, Memory & Cognition, 18*(2), 379-390. doi: 10.1037/0278-7393.18.2.379

Bradley, M.M., & Lang, P.J. (2000). Measuring Emotion: Behavior, Feeling and Physiology. In R.D. Lane & L. Nadel (Eds.), *Cognitive Neuroscience of Emotion* (pp. 242-276). New York: Oxford University Press.

Bradley, M.M., Miccoli, L., Escrig, M.A., & Lang, P.J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology,45*(4), 602-607. doi: 10.1111/j.1469-8986.2008.00654.x

Bradley, M.M., Sabatinelli, D., Lang, P.J., Fitzsimmons, J.R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience, 117*(2), 369-380. doi: 10.1037/0735-7044.117.2.369

Bruce, V., & Young, A. (1986). Understanding Face Recognition. *British Journal of Psychology, 77*(3), 305-327. doi: 10.1111/j.2044-8295.1986.tb.02199.x

Buchanan, T.W. (2007). Retrieval of emotional memories. *Psychological Bulletin, 133*(5), 761-779. doi: 10.1037/0033-2909.133.5.761

Burton, A.M., White, D., & Mcneill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods, 42*(1), 286-91. doi: 10.3758/BRM.42.1.286

Caharel, S., d'Arripe, O., Ramon, M., Jacques, C., & Rossion, B. (2009). Early adaptation to repeated unfamiliar faces across viewpoint changes in the right hemisphere: evidence from the N170 ERP component. *Neuropsychologia, 47*(3), 639-643. doi: 10.1016/j.neuropsychologia.2008.11.016

Calder, A.J., Beaver, J.D., Winston, J.S., Dolan, R.J., Jenkins, R., Eger, E., & Henson, R.N. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology, 17*(1), 20-25. doi: 10.1016/j.cub.2006.10.052

Calder, A.J., Jenkins, R., Cassel, A., & Clifford, C.W. (2008). Visual representation of eye gaze is coded by a nonopponent multichannel system. *Journal of Experimental Psychology: General, 137*(2), 244-261. doi: 10.1037/0096-3445.137.2.244

Campanella, S., Hanoteau, C., Depy, D., Rossion, B., Bruyer, R., Crommelinck, M., & Guerit, J.M. (2000). Right N170 modulation in a face discrimination task: an account for categorical perception of familiar faces. *Psychophysiology, 37*(6), 796-806. doi: 10.1111/1469-8986.3760796

Campbell-Meiklejohn, D., Kanai, R., Bahrami, B., Bach, D., Dolan, R., Roepstorff, A., & Frith, C.D. (2012). Structure of orbitofrontal cortex predicts social influence. *Current Biology, 22*(4), 123-124. doi: 10.1016/j.cub.2012.01.012

Cannon, P.R., Hayes, A.E., & Tipper, S.P. (2009). An electromyographic investigation of the impact of task relevance on facial mimicry. *Cognition & Emotion, 23*(5), 918-929. doi: 10.1080/02699930802234864

Cannon, P.R., Hayes, A.E., & Tipper, S.P. (2010). Sensorimotor fluency influences affect: Evidence from electromyography. *Cognition & Emotion, 24*(4), 681-691. doi: 10.1080/02699930902927698

Cannon, P.R., Schnall, S., & White, M. (2011). Transgressions and expressions: Affective facial muscle activity predicts moral judgments. *Social Psychological and Personality Science, 2*(3), 325-331. doi: 10.1177/1948550610390525

Capozzi, F., Bayliss, A.P., Elena, M.R., & Becchio, C. (2015). One is not enough: Group size modulates social gaze-induced object desirability effects. *Psychonomic Bulletin & Review, 22*(3), 850-855. doi: 10.3758/s13423-014-0717-z

Carlin, J.D., & Calder, A.J. (2013). The neural basis of eye gaze processing. *Current Opinion in Neurobiology, 23,* 450-455. doi: 10.1016/j.conb.2012.11.014

Carlin, J.D., Calder, A.J., Kriegeskorte, N., Nili, H., & Rowe, J.B. (2011). A head view-invariant representation of gaze direction in anterior superior temporal sulcus. *Current Biology, 21*(21), 1817-1821. doi: 10.1016/j.cub.2011.09.025

Caspers, S., Zilles, K., Laird, A.R., & Eickhoff, S.B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage, 50*(3), 1148-1167. doi: 10.1016/j.neuroimage.2009.12.112

Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage, 12*(3), 314-325. doi: 10.1016/nimg.2000.0612

Chang, L.J., Doll, B.B., van't Wout, M., Frank, M.J., & Sanfey, A.G. (2010). Seeing is believing: trustworthiness as a dynamic belief. *Cognitive Psychology, 61*(2), 87-105. doi: 10.1016/j.cogpsych.2010.03.001

Chapman, H.A., Kim, D.A., Susskind, J.M., & Anderson, A.K. (2009). In bad taste: evidence for the oral origins of moral disgust. *Science, 323*(5918), 1222-1228. doi: 10.1126/science.1165565

Cheal, M., & Lyon, D.R. (1991). Central and peripheral precuing of attention of forced-choice discrimination. *Quarterly Journal of Experimental Psychology A, 43*(4), 859-880. doi: 10.1080/14640749108400960

Colby, C.L., & Goldberg, M.E. (1999). Space and Attention in Parietal Cortex. *Annual Review of Neuroscience, 22,* 319-349. doi: 10.1146/annurev.neuro.22.1.319

Collins, J.A., & Olson, I.R. (2014). Beyond the FFA: The role of the ventral anterior temporal lobes in face processing. *Neuropsychologia, 61,* 65-79. doi: 10.1016/j.neuropsychologia.2014.06.005

Cook, I.A., O'Hara, R., Uijdehaage, S.H., Mandelkern, M., & Leuchter, A.F. (1998). Assessing the accuracy of topographic EEG mapping for determining local brain function. *Electroencephalography and Clinical Neuropsysiology, 107*(6), 408-414. doi: 10.1016/S0013-4694(98)00092-3

Corbetta, M., & Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-205. doi: 10.1038/nrn755

Corbetta, M., Patel, G., & Shulman, G.L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron, 58*(3), 306-324. doi: 10.1016/j.neuron.2008.04.017

Craig, A.D. (2003). Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology, 13*(4), 500-505. doi: 10.1016/S0959-4388(03)00090-4

Craig, A.D. (2009). How do you feel now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10*(1), 59-70. doi: 10.1038/nrn2555

Critchley, H.D. (2009). Psychophysiology of neural, cognitive and affective integration: fMRI and autonomic indicants. *International Journal of Psychophysiology, 73,* 88-94. doi: 10.1016/j.ijpsycho.2009.01.012

Cunningham, W.A., Espinet, S.D., DeYoung, C.G., & Zelazo, P.D. (2005). Attitudes to the right- and left: Frontal ERP asymmetries associated with stimulus valence and processing goals. *Neuroimage, 28*(4), 827-834. doi: 10.1016/j.neuroimage.2005.04.044

Cunningham, W.A., & Zelazo, P.D. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. *Trends in Cognitive Neuroscience, 11*(3), 97-104. doi: 10.1016/j.tics.2006.12.005

Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, M., & Lang, P.J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological Psychology, 52*(2), 95-111. doi: 10.1016/S0301-0511(99)00044-7

Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2012). Social status gates social attention in humans. *Biology Letters, 8*(3), 450-452. doi: 10.1098/rsbl.2011.0881

Damasio, A.R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London, B, 351,* 1413-1420. doi: 10.1098/rstb.1996.0125

Damasio, A.R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience, 13,* 89-109. doi: 10.1146/annurev.ne.13.030190.000513

Davey, G.C. (1994). Is evaluative conditioning a qualitatively distinct form of classical conditioning? *Behvaviour Research and Therapy, 32*(3), 291-299. doi: 10.1016/0005-7967(94)90124-4

Davidson, R.J. (1998). Affective style and affective disorders: Perspectives from affective neuroscience. *Cognition and Emotion, 12*(3), 307-330. doi: 10.1080/026999398379628

Davies-Thompson, J., Gouws, A., & Andrews, T.J. (2009). An image-dependent representation of familiar and unfamiliar faces in the human ventral stream. *Neuropsychologia, 47*(6), 1627-1635. doi:10.1016/j.neuropsychologia.2009.01.017

Deaner, R.O., & Platt, M.L. (2003). Reflexive Social Attention in Monkeys and Humans. *Current Biology, 13*, 1609-1613. doi: 10.1016/j.cub.2003.08.025

De Cesarei, A., & Codispoti, M. (2011). Affective modulation of the LPP and α-ERP during picture viewing. *Psychophysiology, 48*(10), 1397-1404. doi: 10.1111/j.1469-8986.2011.01204.x

De Houwer, J. (2006). Using the Implicit Association Test does not rule out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and Motivation, 37,* 176-187. doi: 10.1016/j.lmot.2005.12.002

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: a review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*(6), 853-869. doi: 10.1037/0033-2909.127.6.853

Delgado, M.R., Frank, R.H., & Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience, 8*(11), 1611-1618. doi: 10.1038/nn1575

Demos, K.E., Kelley, W.M., Ryan, S.L., Davis, F.C., & Whalen, P.J. (2008). Human amygdala sensitivity to the pupil size of others. *Cerebral Cortex, 18*(12), 2729-2734. doi: 10.1093/cercor/bhn034

den Ouden, H.E., Frith, U., Frith, C., & Blakemore, S.J. (2005). Thinking about intentions. *Neuroimage, 28*(4), 787-796. doi: 10.1016/j.neuroimage.2005.05.001

Dering, B., Martin, C.D., Moro, S., Pegna, A.J., & Thierry, G. (2011). Face-sensitive processes one hundred milliseconds after picture onset. *Frontiers in Human Neuroscience, 5*(93), 1-14. doi: 10.3389/fnhum.2011.00093

DeSteno, D. (2014). *The truth about trust: How it determines success in life, love, learning, and more.* New York: Penguin.

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology, 19*(6), 643-647. doi: 10.1111/j.1469-8986.1982.tb02516.x

Dimberg, U., & Lundquist, L.O. (1990). Gender differences in facial reactions to facial expressions. *Biological Psychology, 30*(2), 151-159. doi: 10.1016/0301-0511(90)90024-Q

Dimberg, U., & Petterson, M. (2000). Facial reactions to happy and angry facial expressions: evidence for right hemisphere dominance. *Psychophysiology, 37*(5), 693-696. doi: 10.111/1469-8986.3750693

Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology, 39*(1), 39-45. doi: 10.1111/1467-9450.00054

Dimberg, U., & Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expression. *Psychological Science, 11*(1), 86-89. doi: 10.1111/1467-9280.00221

Dimberg, U., Thunberg, M., & Grunedal, S. (2002). Facial reactions to emotional stimuli: Automatically controlled emotional responses. *Cognition & Emotion, 16*(4), 449-471. doi: 10.1080/02699930143000356

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91*(1), 176-180. doi: 10.1007/BF00230027

Dobson, S.D. (2009). Socioecological correlates of facial mobility in nonhuman anthropoids. *American Journal of Physical Anthropology, 139*(3), 413-420. doi: 10.1002/ajpa.21007

Dobson, S.D. (2012). Face to face with the social brain. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 367*(1597), 1901-1908. doi: 10.1098/rstb.2011.0224

Dobson, S.D., & Sherwood, C.C. (2011). Correlated evolution of brain regions involved in producing and processing facial expressions in anthropoid primates. *Biology Letters, 7*(1), 86-88. doi: 10.1098/rsbl.2010.0427

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: encoding pleasant, unpleasant, and neutral pictures. *Cognitive, Affective & Behavioral Neuroscience, 2*(3), 252-263. doi: 10.3758/CABN.2.3.252

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze Perception Triggers Reflexive Visuospatial Orienting. *Visual Cognition, 6*(5), 509-540. doi: 10.1080/135062899394920

Dunn, J.R., & Schweitzer, M.E. (2005). Feeling and Believing: The Influence of Emotion on Trust. *Journal of Personality and Social Psychology, 88*(5), 736-748. doi: 10.1037/0022-3514.88.5.736

Dunning, J.P., & Hajcak, G. (2009). See no evil: directing visual attention within unpleasant images modulates the electrocortical response. *Psychophysiology, 46*(1), 28-33. doi:10.111/j.1469-8986.2008.00723.x

Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia, 45*(1), 15-31. doi: 10.1016/j.neuropsychologia.2006.04.022

Eimer, M., Holmes, A., & McGlone, F.P. (2003). The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. *Cognitive, Affective & Behavioral Neuroscience, 3*(2), 97-110. doi: 10.3758/CABN.3.2.97

Ekman, P., & Friesen, W.V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124-129. doi: 10.1037/h0030377

Ekman, P., & Friesen, WV. (1978). *Facial Action Coding System (FACS): A technique for the measurement of facial action.* Palo Alto, CA: Consulting Psychologists Press.

Ellis, A.W., Young, A.W., & Critchley, E.M. (1989). Loss of memory for people following temporal lobe damage. *Brain, 112*(6), 1469-1483. doi: 10.1093/brain/112.6.1469

Emery, N.J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24,* 581-604. doi: 10.1016/S0149-7634(00)00025-7

Engell, A.D., Haxby, J.V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience, 19*(9), 1508-1519. doi: 10.1162/jocn.2007.19.9.1508

Engell, A.D., Todorov, A., & Haxby, J.V. (2010). Common neural mechanisms for the evaluation of facial trustworthiness and emotional expressions as revealed by behavioural adaptation. *Perception, 39*(7), 931-941. doi: 10.1068/p6633

Evans, J.J., Heggs, A.J., Antoun, N., & Hodges, J.R. (1995). Progressive prosopagnosia associated with selective right temporal lobe atrophy. A new syndrome? *Brain, 118*(1), 1-13. doi: 10.1093/brain/118.1.1

Farroni, T., Massaccesi, S., Pividori, D., Simion, F., & Johnson, M.H. (2004). Gaze Following in Newborns. *Infancy, 5*(1), 39-60. doi: 10.1207/s15327078in0501_2

Fehr, E., & Camerer, C.F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences, 11*(10), 419-427. doi: 10.1016/j.tics.2007.09.002

Feinstein, J.S., Adolphs, R., Damasio, A., & Tranel, D. (2011). The human amygdala and the induction and experience of fear. *Current Biology, 21*(1), 34-38. doi: 10.1016/j.cub.2010.11.042

Ferri, J., Weinberg, A., & Hajcak, G. (2012). I see people: The presence of human faces impacts the processing of complex emotional stimuli. *Social Neuroscience, 7*(4), 436-443. doi: 10.1080/17470919.2012.680492

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe from action organization to intention understanding. *Science, 308*(5722), 662-667. doi: 10.1126/science.1106138

Foti, D., & Hajcak, G. (2008). Deconstructing reappraisal: descriptions preceding arousing pictures modulate the subsequent neural response. *Journal of Cognitive Neuroscience, 20*(6), 977-988. doi: 10.1162/jocn.2008.20066

Franken, I.H., Huijding, J., Nijs, I.M., & van Strein, J.W. (2011). Electrophysiology of appetitive taste and appetitive taste conditioning in humans. *Biological Psychology, 86*(3), 273-278. doi: 10.1016/j.biopsycho.2010.12.008

Freeman, J.B., Stolier, R.M., Ingbretsen, Z.A., & Hehman, E.A. (2014). Amygdala responsivity to high-level social information from unseen faces. *Journal of Neuroscience, 34*(32), 10573-10581. doi: 10.1523/JNEUROSCI.5063-13.2014

Freese, J.L., & Amaral, D.G. (2005). The organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. *Journal of Comparative Neurology, 486*(4), 295-317. doi: 10.1002/cne.20520

Freese, J.L., & Amaral, D.G. (2009). Neuroanatomy of the primate amygdala. In P.J. Whalen, & E.A. Phelps (Eds.), *The Human Amygdala* (pp.3-42). New York: Guilford Press.

Freiwald, W.A., & Tsao, D.Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science, 330*(6005), 845-851. doi: 10.1126/science.1194908

Fridlund, A.J., & Cacioppo, J.T. (1986). Guidelines for human electromyographic research. *Psychophysiology, 23*(5), 567-589. 10.1111/j.1469-8986.1986.tb00676.x

Friesen, C.K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review, 5*(3), 490-495. doi: 10.3758/BF03208827

Frijda, N.H. (1986). *The Emotions*. Cambridge UK: Cambridge University Press.

Frisby, J.P. (1979). *Seeing: Illusion, brain and mind.* Oxford, England: Oxford University Press.

Frischen, A., Bayliss, A.P., & Tipper, S.P. (2007). Gaze-cueing of attention: Visual attention, social cognition and individual differences. *Psychological Bulletin, 133*(4), 694-724. doi: 10.1037/0033-2909.133.4.694

Friston, K.J., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., & Mattout,J. (2008). Multiple sparse priors for the M/EEG inverse problem. *Neuroimage, 39*(3), 1104-1120. doi: 10.1016/j.neuroimage.2007.09.048

Gainotti, G. (1972). Emotional behaviour and hemispheric side of the lesion. *Cortex, 8,* 41-55. doi: 10.1016/S0010-9452(72)80026-1

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(2), 593-609. doi: 10.1093/brain/119.2.593

Gasparrini, W.G., Satz, P., Heilman, K., & Coolidge, F.L. (1978). Hemispheric asymmetries of affective processingas determined by the Minnesota Multiphasic Personality Inventory. *Journal of Neurology, Neurosurgery & Psychiatry, 41*(5), 470-473. doi: 10.1136/jnnp.41.5.470

Gibson, J., & Pick, A. (1963). Perception of another persons looking behavior. *American Journal of Psychology, 76*(3), 386-394. doi: 10.2307/1419779

Glenberg, A.M., & Kaschak, M.P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*(3), 558-565. doi: 10.3758/BF03196313

Gonzalez, J., Barros-Loscertales, A., Pulvermuller, F., Meseguer, V., Sanjuan, A., Belloch, V., & Avila, C. (2006). Reading cinnamon activates olfactory brain regions. *Neuroimage, 32*(2), 906-912. doi: 10.1016/j.neuroimage.2006.03.037

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

Gothard, K. M. (2014). The amygdalo-motor pathways and the control of facial expressions. *Frontiers in Neuroscience, 8*(43), 1-7. doi: 10.3389/fnins.2014.00043

Gothard, K.M., Battaglia, F.P., Erickson, C.A., Spitler, K.M., & Amaral, D.G. (2007). Neural responses to facial expression and face identity in the monkey amygdala. *Journal of Neurophysiology, 97*(2), 1671-1683. doi: 10.1152/jn.00714.2006

Greenwald, M.K., Cook, E.W., & Lang, P.J. (1989). Affective judgment and psychophysiological response: dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology, 3,* 51-64.

Guillaume, F., & Tiberghien, G. (2001). An event-related potential study of contextual modifications in a face recognition task. *Neuroreport, 12*(6), 1209-1216.

Gurtman, M.B., & Lion, C. (1982). Interpersonal trust and perceptual vigilance for trustworthiness descriptors. *Journal of Research in Personality, 16*(1), 108-117. doi: 10.1016/0092-6566(82)90044-7

Hajcak, G., Dunning, J.P., & Foti, D. (2007). Neural response to emotional pictures is unaffected by concurrent task difficulty: an event-related potential study. *Behavioral Neuroscience, 121*(6), 1156-1162. doi: 10.1037/0735-7044.121.6.1156

Hajcak, G., Dunning, J.P., & Foti, D. (2009). Motivated and controlled attention to emotion: time-course of the late positive potential. *Clinical Neurophysiology, 120*(3), 505-510. doi: 10.1016/j.clinph.2008.11.028

Hajcak, G., Macnamara, A., & Olvet, D.M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology, 35*(2), 129-155. doi: 10.1080/87565640903526504

Hajcak, G., & Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cognitive, Affective and Behavioral Neuroscience, 6*(4), 291-297. doi: 10.3758/CABN.6.4.291

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the Study of Emotion. In S.J. Luck & E.S. Kappenman, *The Oxford Handbook of Event-Related Potential Components* (pp.441-472). New York: Oxford University Press.

Hareli, S., & Parkinson, B. (2008). What's social about social emotions? *Journal for the Theory of Social Behaviour, 38*(2), 131-156. doi: 10.1111/j.1468-5914.2008.00363.x

Harmon-Jones, E. (2006). Unilateral right-hand contractions cause contralateral alpha power suppression and approach motivational affective experience. *Psychophysiology, 43,* 598-603. doi: 10.1111/j.1469-8986.2006.00465.x

Harmon-Jones, E., Gable, P.A., & Peterson, C.K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biological Psychology, 84,* 451-462. doi: 10.1016/j.biopsycho.2009.08.010

Harris, A., Adolphs, R., Camerer, C., & Rangel, A. (2011). Dynamic construction of stimulus values in the ventromedial prefrontal cortex. *Plos One, 6*(6), 1-7. doi: 10.1371/journal.pone.0021074

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*(2), 301-307. doi: 10.1016/S0896-6273(03)00838-9

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. doi: 10.1016/S1364-6613(00)01482-0

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. doi: 10.1016/S0006-3223(01)01330-0

Hebb, D.O. (1949). *The organization of behaviour. A neuropsychological theory*. New York: Wiley.

Heller, A.S., Greischar, L.L., Honor, A., Anderle, M.J., & Davidson, R.J. (2011). Simultaneous acquisition of corrugator electromyography and functional magnetic resonance imaging: a new method for objectively measuring affect and neural activity concurrently. *Neuroimage, 58*(3), 930-934. doi: 10.1016/j.neuroimage.2011.06.057

Heller, A.S., Lapate, R.C., Mayer, K.E., & Davidson, R.J. (2014). The face of negative affect: trial-by-trial corrugator responses to negative pictures are positively associated with amygdala and negatively associated with ventromedial prefrontal cortex activity. *Journal of Cognitive Neuroscience, 26*(9), 2102-2110. doi: 10.1162/jocn_a_00622

Henson, R.N., Mouchlianitis, E., & Friston, K.J. (2009). MEG and EEG data fusion: Simultaneous localisation of face-evoked responses. *Neuroimage, 47*(2), 581-589. doi: 10.1016/j.neuroimage.2009.04.063

Herrmann, M.J., Ehlis, A-C., Ellgring, H., & Fallgatter, A.J. (2005). Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of Neural Transmission, 112*(8), 1073-1081. doi: 10.1007/s00702-004-0250-8

Hilgard, J., Weinberg, A., Proudfit, G., Bartholow, B.D. (2014). The negativity bias in affective picture processing depends on top-down and bottom-up motivational significance. *Emotion, 14*(5), 940-949. doi: 10.1037/a0036791

Hofman, D., Bos, P.A., Schutter, D.J., & van Honk, J. (2012). Fairness modulates non-conscious facial mimicry in women. *Proceedings of the Royal Society B: Biological Sciences, 279*(1742), 3535-3539. doi: 10.1098/rspb.2012.0694

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychological Bulletin, 136*(3), 390-421. doi: 10.1037/a0018916

Holmes, A., Winston, J.S., & Eimer, M. (2005). The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. *Cognitive Brain Research, 25*(2), 508-520. doi: 10.1016/j.cogbrainres.2005.08.003

Hood, B.M., Willen, J.D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science, 9,* 131-134. doi: 10.1111/1467-9280.00024

Hopf, H.C., Muller-Forell, W., & Hopf, N.J. (1992). Localization of emotional and volitional facial paresis. *Neurology, 42,* 1918-1923. doi: 10.1212/WNL.42.10.1918

Hornak, J., Bramham, J., Rolls, E.T., Morris, R.G., O'Doherty, J., Bullock, P.R., & Polkey, C.E. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain, 126*(7), 1691-1712. doi: 10.1093/brain/awg168

Huddy, V., Schweinberger, S.R., Jentzsch, I., & Burton, M. (2003). Matching faces for semantic information and names: an event-related brain potentials study. *Cognitive Brain Research, 17*(2), 314-326. doi: 10.1016/S0926-6410(03)00131-9

Hudson, M., Nijboer, T.C., & Jellema, T. (2012). Implicit social learning in relation to autistic-like traits. *Journal of Autism and Developmental Disorders, 42*(12), 2534-2545. doi: 10.1007/s10803-012-1510-3

Huijgen, J., Dinkelacker, V., Lachat, F., Yahia-Cherif, L., Karoui, I.E., Lemarechal, J.-D., et al. (2015). Amygdala processing of social cues from faces: an intercerebral EEG study. *Social Cognitive & Affective Neuroscience.* doi: 10.1093/scan/nsv048

Ibanez, A., Melloni, M., Huepe, D., Helgiu, E., Rivera-Rei, A., Canales-Johnson, A., Baker, P., & Moya, A. (2012). What event-related potentials (ERPs) bring to social neuroscience? *Social Neuroscience, 7*(6), 632-649. doi: 10.1080/17470919.2012.691078

Itier, R.J., Latinus, M., & Taylor, M.J. (2006). Face, eye and object early processing: what is the face specificity? *Neuroimage, 29*(2), 667-676. doi: 10.1016/j.neuroimage.2005.07.041

Itier, R.J., & Taylor, M.J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *Neuroimage, 15*(2), 353-372. doi: 10.1006/nimg.2001.0982

Ito, T.A., Larsen, J.T., Smith, N.K., & Cacioppo, J.T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorisations. *Journal of Personality and Social Psychology, 75*(4), 887-900. doi: 10.1037/0022-3514.75.4.887

Izard, C.E. (1993). Four systems for emotion activation: cognitive and noncognitive processes. *Psychological Review, 100*(1), 68-90. doi: 10.1037/0033295X.100.1.68

Jacques, C., d'Arripe, O., & Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of Vision, 7*(3), 1-9. doi: 10.1167/7.8.3

James, W. (1884). What is an emotion? *Mind, 9,* 188-205. doi: 10.1093/mind/os-IX.34.188

Jeffereys, D.A. (1996). Evoked potential studies of face and object processing. *Visual Cognition, 3*(1), 1-38. doi: 10.1080/713756729

Jenkins, R., Beaver, J.D., & Calder, A.J. (2006). I thought you were looking at me: direction-specific aftereffects in gaze perception. *Psychological Science, 17*(6), 506-513. doi: 10.1111/j.1467-9280.2006.01736.x

Johnson-George, C., & Swap, W.C. (1982). Measurement of specific interpersonal trust: construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology, 43*(6), 1306-1317. doi: 10.1037/00223514.43.6.1306

Jones, C.R., Fazio, R.H., & Olson, M.A. (2009). Implicit misattribution as a mechanism underlying evaluative conditioning. *Journal of Personality & Social Psychology, 96*(5), 933-948. doi: 10.1037/a0014747

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. doi:10.1016/j.clinph.2005.07.005

Kanwisher, N., & Barton, J.J.S. (2011). The functional architecture of the face system: integrating evidence from fMRI and patient studies. In A.J. Calder, G. Rhodes, M.H. Johnson & J.V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp, 111-129). New York: Oxford University Press.

Kanwisher, N., McDermott, J., & Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302-4311.

Karnosh, L.J. (1945). Amimia or emotional paralysis of the face. *Diseases of the Nervous System, 6,* 106-108.

Keil, A., Bradley, M.M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P.J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology, 39*(5), 641-649. doi: 10.1111/1469-8986.3950641

Keysers, C., Wicker, B., Gazzola, V., Anton, J.L., Fogassi, L., & Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron, 42*(2), 335-346. doi: 10.1016/S0896-6273(04)00156-4

Kiefer, M., Sim, E.J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. *Journal of Neuroscience, 28*(47), 12224-12230. doi: 10.1523/JNEUROSCI.3579-08.2008

Kilner, J.M. (2013). Bias in a common EEG and MEG statistical analysis and how to avoid it. *Clinical Neurophysiology,124*(10), 2062-2063. doi: 10.1016/j.clinph.2013.03.024

Kilner, J.M., & Friston, K.J. (2010). Topological inference for EEG and MEG. *The Annals of Applied Statistics, 4*(3), 1272-1290. doi: 10.1214/10-AOAS337

Kim, H., Choi, M.J., & Jang, I.J. (2012). Lateral OFC activity predicts decision bias due to first impressions during ultimatum games. *Journal of Cognitive Neuroscience, 24*(2), 428-439. doi: 10.1162/jocn_a_00136

King-Casas, B., Tomlin, D., Anen, C., Camerer, C.F., Quartz, S.R., & Montague, P.R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science, 308*(5718), 78-83. doi: 10.1126/science.1108062

Klein, J.T., Deaner, R.O., & Platt, M.L. (2008). Neural correlates of social target value in macaque parietal cortex. *Current Biology, 18*(6), 419-424. doi: 10.1016/j.cub.2008.02.047

Klein, J.T., Shepherd, S.V., & Platt, M.L. (2009). Social Attention and the Brain. *Current Biology, 9,* R958-R962. doi: 10.1016/j.cub.2009.08.010

Kloth, N., Schweinberger, S.R., & Kovacs, G. (2010). Neural correlates of generic versus gender-specific face adaptation. *Journal of Cognitive Neuroscience, 22*(10), 2345-2356. doi: 10.1162/jocn.2009.21329

Kobayashi, H., & Kohshima, S. (1997). Unique Morphology of the Human Eye. *Nature, 387*(6635), 767-768. doi: 10.1038/42842

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,* 419-435. doi: 10.1006/jhev.2001.0468

Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science, 297*(5582), 846-848. doi: 10.1126/science.1070311

Kohls, G., Perino, M.T., Taylor, J.M., Madva, E.N., Cayless, S.J., Troiani, V., et al. (2013). The nucleus accumbens is involved in both the pursuit of social reward and the avoidance of social punishment. *Neuropsychologia, 51,* 2062-2069. doi: 10.1016/j.neuropsychologia.2013.07.020

Konorski, J. (1967). *Integrative activity of the brain: An interdisciplinary approach.* Chicago: University of Chicago Press.

Koscik, T.R., & Tranel, D. (2011). The human amygdala is necessary for developing and expressing normal interpersonal trust. *Neuropsychologia, 49*(4), 601-611. doi: 10.1016/j.neuropsychologia.2010.09.023

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences USA,104*(51), 20600-20605. doi: 10.1073/pnas.0705654104

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S., & Baker, C.I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience, 12*(5), 535-540. doi: 10.1038/nn.2303

Krumhuber, E., Manstead, A.S.R., Cosker, D., Marshall, D., Rosin, P.L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behaviour. *Emotion, 7,* 730-735. doi: 10.1037/15283542.7.4.730

Kuhn, G., & Tipples, J. (2011). Increased gaze following for fearful faces. It depends on what you're looking for! *Psychonomic Bulletin & Review, 18*(1), 89-95. doi: 10.3758/s13423-010-0033-1

Kunde, W., Skirde, S., & Weigelt, M. (2011). Trust My Face: Cognitive Factors of Head Fakes in Sports. *Journal of Experimental Psychology: Applied, 17*(2), 110-127. doi: 10.1037/a0023756

Laane, K., Aru, J., & Dickinson, A. (2010). Non-competitive liking for brands. No blocking in evaluative conditioning. *Appetite, 54,* 100-107. doi: 10.1016/j.appet.2009.09.012

LaBar, K.S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*(1), 54-64. doi: 10.1038/nrn1825

LaBar, K.S., & Phelps, E.A. (1998). Arousal-Mediated Memory Consolidation: Role of the Medial Temporal Lobe in Humans. *Psychological Science, 9*(6), 490-493. doi: 10.1111/1467-9280.00090

Lachat, F., Conty, L., Hugueville, L., & George, N. (2012a). Gaze cueing effect in a face-to-face situation. *Journal of Nonverbal Behavior, 36,* 177-190. doi: 10.1007/s10919-012-0133-x

Lachat, F., Hugueville, L., Lemarechal, J.D., Conty, L., & George, N. (2012b). Oscillatory brain correlates of live joint-attention: A dual-EEG study. *Frontiers in Human Neuroscience, 6*(156), 1-12. doi: 10.3389/fnhum.2012.00156

Lang, P.J., Greenwald, M.K., Bradley, M.M., & Hamm, A.O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30,* 261-273. doi: 10.1111/j.1469-8986.1993.tb03352.x

Lange, C. (1885). The emotions. In E. Dunlap (Ed.), *The Emotions*. Baltimore: Williams and Wilkins.

Langton, S.R., O'Donnell, C., Riby, D.M., & Ballantyne, C.J. (2006). Gaze cues influence allocation of attention in natural scene viewing. *Quarterly Journal of Experimental Psychology, 59*(12), 2056-2064. doi: 10.1080/17470210600917884

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), 50-58. doi: 10.1016/S1364-6613(99)01436-9

Larsen, J.T., & Norris, J.I. (2009). A facial electromyographic investigation of affective contrast. *Psychophysiology, 46*(4), 831-842. doi: 10.1111/j.1469-8986.2009.00820.x

Larsen, J.T., Norris, C.J., & Cacioppo, J.T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology, 40,* 776-785. doi: 10.1111/1469-8986.00078

Larsen, T., & O'Doherty, J.P. (2014). Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: a combined fMRI-EEG study. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 369*(1655), 1-10. doi: 10.1098/rstb.2013.0473

Law, A.S., Langton, S.R.H., & Logie, R.H. (2010). Assessing the impact of verbal and visuospatial working memory load on eye-gaze cueing. *Visual Cognition, 18*(10), 1420,1438. doi: 10.1080/13506285.2010.496579

Lazarus, R.S. (1991). *Emotion and Adaptation.* New York: Oxford University Press.

Ledoux, J.E. (2000). Emotion circuits in the brain. *Annual Reviews of Neuroscience, 23,* 155-184. doi: 10.1146/annurev.neuro.23.1.155

Leonard, C.M., Rolls, E.T., Wilson, F.A., & Baylis, G.C. (1985). Neurons in the amygdala of the monkey with responses selective for faces. *Behavioural and Brain Research, 15*(2), 159-176. doi: 10.1016/0166-4328(85)90062-2

Leopold, D.A., Bondar, I.V., & Giese, M.A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature, 442*(7102), 572-575. doi: 10.1038/nature04951

Leopold, D.A., O'Toole, A.J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*(1), 89-94. doi: 10.1038/82947

Letourneau, S.M., & Mitchell, T.V. (2008). Behavioral and ERP measures of holistic face processing in a composite task. *Brain & Cognition, 67*(2), 234-245. doi: 10.1016/j.bandc.2008.01.007

Levita, L., Hare, T., Voss, H.U., Glover, G., Ballon, D.J., & Casey, B.J. (2009). The bivalent side of the nucleus accumbens. *Neuroimage, 44*(3), 1178-1187. doi: 10.1016//j.neuroimage.2008.09.039

Levita, L., Hoskin, R., & Champi, S. (2012). Avoidance of harm and anxiety: a role for the nucleus acccumbens. *Neuroimage, 62*(1), 189-198. doi: 10.1016/j.neuroimage.2012.04.059

Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social, Cognitive & Affective Neuroscience, 7*(3), 274-281. doi: 10.1093/scan/nsr006

Litvak, V., & Friston, K. (2008). Electromagnetic source reconstruction for group studies. *Neuroimage, 42*(4), 1490-1498. doi: 10.1016/j.neuroimage.2008.06.022

Litvak,V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., Barnes, G., et al. (2011). EEG and MEG data analysis in SPM8. *Computational Intelligence and Neuroscience, 852961,*1-32. doi: 10.1155/2011/852961

Liu, J., Harris, A., & Kanwisher, N. (2009). Perception of face parts and face configurations: An fMRI study. *Journal of Cognitive Neuroscience, 22*(1), 203-211. doi: 10.1162/jocn.2009.21203

Liu, Y., Huang, H., McGinnis-Deweese, M., Keil, A., & Ding, M. (2012). Neural substrate of the late positive potential in emotion processing. *The Journal of Neuroscience, 32*(42), 14563-14572. doi: 10.1523/JNEUROSCI.3109-12.2012

Livneh, U., Resnik, J., Shohat, Y., & Paz, R. (2012). Self-monitoring of social facial expressions in the primate amygdala and cingulate cortex. *Proceedings of the National Academy of Sciences USA, 109,* 18956-18961. doi: 10.1073/pnas.1207662109

Lobmaier, J.S., Tiddeman, B.P., & Perrett, D.I. (2008). Emotional expression modulates perceived gaze direction. *Emotion, 8*(4), 573-577. doi: 10.1037/1528-3542.8.4.573

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

Machado, C.J., Emery, N.J., Capitanio, J.P., Mason, W.A., Mendoza, S.P., & Amaral, D.G. (2008). Bilateral neurotoxic amygdala lesions in rhesus monkeys (Macaca mulata): consistent pattern of behaviour across different social contexts. *Behavioral Neuroscience, 122*(2), 251-266. doi: 10.1037/0735-7044.122.2.251

Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276-298. doi: 10.1037/h0076778

MacNamara, A., Ferri, J., & Hajcak, G. (2011). Working memory load reduces the late positive potential and this effect is attenuated with increasing anxiety. *Cognitive, Affective & Behavioral Neuroscience, 11*(3), 321-331. doi: 10.3758/s13415-011-0036-z

MacNamara, A., Foti, D., & Hajcak, G. (2009). Tell me about it: neural activity elicited by emotional pictures and preceding descriptions. *Emotion, 9*(4), 531-43. doi: 10.1037/a0016251

Magnee, M.J., Stekelenburg, J.J., Kemner, C., & de Gelder, B. (2007). Similar facial electromyographic responses to faces, voices, and body expressions. *Neuroreport, 18*(4), 369-372. doi: 10.1097/WNR.0b013e32801776e6

Manera, V., Elena, M.R., Bayliss, A.P., & Becchio, C. (2014). When seeing is more than looking: Intentional gaze modulates object desirability. *Emotion, 14*(4), 824-832. doi: 10.1037/a0036258

Martin, I., & Levey, A. (1978). Evaluative Conditioning. *Advances in Behaviour Research and Therapy, 1,* 57-102. doi: 10.1016/0146-6402(78)90013-9

Martin, I., & Levey, A. (1994). The Evaluative Response: Primitive But Neccessary. *Behavior Research and Therapy, 32(3),* 301-305. doi: 10.1016/0005-7967(94)90125-2

Marzi,T., Righi, S., Ottonello, S., Cincotta, M., & Viggiano, P.M. (2011). Trust at first sight: Evidence from ERPs. *Social Cognitive and Affective Neuroscience, 9*(1), 63-72. doi: 10.1093/scan/nss102

Mattarozzi, K., Todorov, A., & Codispoti, M. (2015). Memory for faces: the effect of facial appearance and the context in which the face is encountered. *Psychological Research, 79*(2), 308-317. doi: 10.1007/s00426-014-0554-8

Mattavelli, G., Andrews, T.J., Asghar, A.U., Towler, J.R., & Young, A.W. (2012). Response of face-selective brain regions to trustworthiness and gender of faces. *Neuropsychologia, 50*(9), 2205-2211. doi: 10.1016/j.neuropsychologia.2012.05.024

Mattavelli, G., Sormaz, M., Flack, T., Asghar, A.U.R., Fan, S., Frey, J., Manssuer, L., Usten, D., Young, A.W., & Andrews T.J. (2013). Neural responses to facial expressions support the role of the amygdala in processing threat. *Social Cognitive and Affective Neuroscience, 9*(11), 1684-1689. doi: 10.1093/scan/nst162

Maurer, U., Rossion, B., & McCandliss, B.D. (2008). Category specificity in early perception: face and word n170 responses differ in both lateralization and habituation properties. *Frontiers in Human Neuroscience, 2*(18), 1-7. doi: 10.3389/neuro.09.018.2008

McCaugh, J.L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Reviews in Neuroscience, 27,* 1-28. doi: 10.1146/annurev.neuro.27.070203.144157

McCoy, A.N., & Platt, M.L. (2005). Expectations and outcomes: decision-making in the primate brain. *Journal of Comparative Physiology A, 191*(3), 201-211. doi: 10.1007/s00359-004-0565-9

Mehu, M., Grammer, K., & Dunbar, R.I.M. (2007). Smiles when sharing. *Evolution & Human Behavior, 28*(6), 415-422. doi: 10.1016/j.evolhumbehav.2007.05.010

Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2013). The neural dynamics of updating person impressions. *Social, Cognitive & Affective Neuroscience, 8*(6), 623-631. doi: 10.1093/scan/nss040

Miltner, W.H., Braun, C.H., & Coles, M.G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience, 9*(6), 788-798. doi: 10.1162/jocn.1997.9.6.788

Moran, T.P., Jendrusina, A.A., & Moser, J.S. (2013). The psychometric properties of the late positive potential during emotion processing and regulation. *Brain Research, 1516,* 66-75. doi: 10.1016/j.brainres.2013.04.018

Morin, E.L., Hadj-Bouziane, F., Stokes, M., Ungerleider, L.G., & Bell, A.H. (2014). Hierarchical Encoding of Social Cues in Primate Inferior Temporal Cortex. *Cerebral Cortex.* doi: 10.1093/cercor/bhu099

Morrati, S., Saugar, C., Strange, B.A. (2011). Prefrontal-Occipitoparietal coupling underlies late latency human neuronal responses to emotion. *The Journal of Neuroscience, 31*(47), 17278-17286. doi: 10.1523/JNEUROSCI.2917-11.2011

Mosher, C.P., Zimmerman, P.E., & Gothard, K.M. (2014). Neurons in the monkey amygdala detect eye contact during naturalistic social interactions. *Current Biology, 24*(20), 2459-2464. doi: 10.1016/j.cub.2014.08.063

Müller, H.J., & Krummenacher, J. (2006). Visual search and selective attention. *Visual Cognition, 14*(4-8), 389-410. doi: 10.1080/13506280500527676

Nagai, Y., Critchley, H.D., Featherstone, E., Trimble, M.R., & Dolan, R.J. (2004). Activity in ventromedial prefrontal cortex covaries with sympathetic skin conductance level: a physiological account of a "default mode" of brain function. *Neuroimage, 22*(1), 243-251. doi: 10.1016/j.neuroimage.2004.01.019

Nasr, S., & Tootell, R.B. (2012). Role of fusiform and anterior temporal cortical areas in face recognition. *Neuroimage, 63*(3), 1743-1753. doi: 10.1016/j.neuroimage.2012.08.031

Neta, M., Norris, C.J., & Whalen, P.J. (2009). Corrugator Muscle Responses Are Associated With Individual Differences in Positivity-Negativity Bias. *Emotion, 9*(5), 640-648. doi: 10.1037/a0016819

Neumann, R., Schulz, S.M., Lozo, L., & Alpers, G.W. (2014). Automatic facial responses to near-threshold presented facial displays of emotion: imitation of evaluation? *Biological Psychology, 96,* 10.1016/j.biopsycho.2013.12.009

Niedenthal, P.M. (2007). Embodying Emotion. *Science, 316*(5827), 1002-1005. doi: 10.1126/science.1136930

O'Doherty, J.P., Buchanan, T.W., Seymour,B., & Dolan, R.J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron, 49*(1), 157-166. doi: 10.1016/j.neuron.2005.11.014

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4*(1), 95-102. doi: 10.1038/82959

O'Doherty, J., Winston, J., Critchley, H., Perrett, D.I., Burt, D.M., & Dolan, R.J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41*(2), 147-155. doi: 10.1016/S0028-3932(02)00145-8

Oosterhof, N.N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences USA, 105*(32), 11087-11092. doi: 10.1073/pnas.0805664105

Osinsky, R., Mussel, P., Ohrlein, L., & Hewig, J. (2014). A neural signature of the creation of social evaluation. *Social, Cognitive & Affective Neuroscience, 9*(6), 731-736. doi: 10.1093/scan/nst051

Over, H., & Carpenter, M. (2009). Priming third-party ostracism increases affiliative imitation in children. *Developmental Science, 12*(3), F1-8. doi: 10.1111/j.1467-7687.2008.00820.x

Palomba, D., Angrilli, A., & Mini, A. (1997). Visual evoked potentials, heart rate responses and memory to emotional pictorial stimuli. *International Journal of Psychophysiology, 27*(1), 55-67. doi: 10.1016/S0167-8760(97)00751-4

Paton, J.J., Belova, M.A., Morrison, S.E., & Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature, 439*(7078), 865-870. doi: 10.1038/nature04490

Pavlov, I.P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex.* Oxford: Oxford University Press.

Pearce, J.M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87*(6), 532-552. doi: 10.1037/0033-295X.87.6.532

Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14*(2), 119-124. doi: 10.1111/1467-9280.t01-1-01429

Pelphrey, K.A., Morris, J.P., & McCarthy, G. (2005). Neural basis of eye gaze processing deficits in autism. *Brain, 128*(5), 1038-1048. doi: 10.1093/brain/awh404

Pelphrey, K.A., Singerman, J.D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia, 41*(2), 156-170. doi: 10.1016/S0028-3932(02)00146-X

Perrett, D.I., Hietanen, J.K., Oram, M.W., & Benson, P.J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences, 335*(1273), 23-30. doi: 10.1098/rstb.1992.0003

Petrican, R., English, T., Gross, J.J., Grady, C., Hai, T., & Moscovitch, M. (2013). Friend or foe? Age moderates time-course specific responsiveness to trustworthiness cues. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences, 68*(2), 215-223. doi: 10.1093/geronb/gbs064

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. doi: 10.1016/j.neuroimage.2014.06.061

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. doi: 10.1016/j.neubiorev.2013.07.017

Pfurtscheller, G., & Lopes da Silva, F.H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology, 110*(11), 1842-1857. doi: 10.1016/S1388-2457(99)00141-8

Pierce, L.J., Scott, L.S., Boddington, S., Droucker, D., Curran, T., & Tanaka, J.W. (2011). The n250 brain potential to personally familiar and newly learned faces and objects. *Frontiers in Human Neuroscience,5,* 111. doi: 10.3389/fnhum.2011.00111

Poli, S., Sarlo, M., Bortoletto, M., Buodo, G., & Palomba, D. (2007). Stimulus-preceding negativity and heart rate changes in anticipation of affective pictures. *International Journal of Psychophysiology, 65*(1), 32-39. doi: 10.1016/j.ijpsycho.2007.02.008

Polich, J. (2012). Neuropsychology of P300. In S.J. Luck & E.S. Kappenman, *The Oxford Handbook of Event-Related Potential Components* (pp.159-188). New York: Oxford University Press.

Pönkänen, L.M., & Hietanen, J.K. (2012). Eye contact with neutral and smiling faces: effects on autonomic responses and frontal EEG asymmetry. *Frontiers in Human Neuroscience, 6*(122), 1-9. doi: 10.3389/fnhum.2012.00122

Porges, S.W. (1997). Emotion: An evolutionary by-product of the neural regulation of the autonomic nervous system. *Annals of the New York Academy of Sciences, 807,* 62-77. doi: 10.1111/j.1749-6632.1997.tb51913.x

Posner, M.I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology, 32* (1), 3-25. doi: 10.1080/00335558008248231

Posner, M.I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531-556).

Poulton, E.C. (1982). Influential companions: Effects of one strategy on another in the within-subjects designs of cognitive psychology. *Psychological Bulletin, 91,* 673-690. doi: 10.1037/0033-2909.91.3.673

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*(7), 576-582. doi: 10.1038/nrn1706

Rajimehr, R., Young, J.C., & Tootell, R.B. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences U.S.A., 106*(6), 1995-2000. doi: 10.1073/pnas.0807304106

Ramsey, R., Cross, E.S., & Hamilton, A.F. de C. (2011). Eye can see what you want: posterior intraparietal sulcus encodes the object of an actor's gaze. *Journal of Cognitive Neuroscience, 23*(11), 3400-3409. doi: 10.1162/jocn_a_00074

Regenberg, N.F., Hafner, M., & Semin, G.R. (2012). The groove move: action affordances produce fluency and positive affect. *Experimental Psychology, 59*(1), 30-37. doi: 10.1027/1618-3169/a000122

Rehmert, A.E., & Kisley, M.A. (2013). Can older adults resist the positivity effect in neural responding? The impact of verbal framing on event-related brain potentials elicited by emotional images. *Emotion, 13*(5), 949-959. doi: 10.1037/a0032771

Rescorla, R.A. & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement, In A.H. Black & W.F. Prokasy (Eds.), *Classical Conditioning II* : Current research and theory (pp. 64–99). New York: Appleton-Century-Crofts.

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research, 46*(18), 2977-2987. doi: 10.1016/j.visres.2006.03.002

Rhodes, G., Michie, P.T., Hughes, M.E., & Byatt, G. (2009). The fusiform face area and occipital face area show sensitivity to spatial relations in faces. *European Journal of Neuroscience, 30*(4), 721-733. doi: 10.1111/j.1460-9568.2009.06861.x

Rhodes, G., Morley, G., & Simmons, L.W. (2012). Women can judge sexual unfaithful from unfamiliar men's faces. *Biology Letters, 9*(1), 1-6. doi: 10.1098/rsbl.2012.0908

Ristic, J., Friesen, C.K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review, 9*(3), 507-513. doi: 10.3758/BF03196306

Robinson, R.G., & Price, T.R. (1982). Post-stroke depressive disorders: a follow-up of 103 patients. *Stroke, 13*(5), 635-641. doi: 10.1161/01.STR.13.5.635

Robinson, J.P., Shaver, P.R., & Wrightsman, L.S. (1991). *Measures of personality and social psychological attitudes.* San Diego: Academic Press.

Rogers, R.D., Bayliss, A.P., Szepietowska, A., Dale, L. Czarna, K., Wakeley, J. & Tipper, S.P. (2014). I want to help you, but I am not sure why: Gaze-cuing induces altruistic giving. *Journal of Experimental Psychology: General, 143*(2), 763-777. doi: 10.1037/a0033677

Rogers, R.D., Bayliss, A.P., Wakeley, J., Cowen, P.J., & Tipper, S.P. (unpublished manuscript). Vulnerability to depression is associated with a failure to acquire implicit social appraisals.

Rolls, E.T. (2008). *Emotion Explained.* New York: Oxford University Press.

Rolls, E.T., Critchley, H.D., Browning, A.S., & Inoue, K. (2006). Face-selective and auditory neurons in the primate orbitofrontal cortex. *Experimental Brain Research, 170*(1), 74-87. doi: 10.1007/s00221-005-0191-y

Rolls, E.T., & Grabenhorst, F. (2008). The orbitofrontal cortex and beyond: from affect to decision-making. *Progress in Neurobiology, 86*(3), 216-244. doi: 10.1016/j.pneurobio.2008.09.001

Roozendaal, B., & McGaugh, J.L. (2011). Memory modulation. *Behavioral Neuroscience, 125*(6), 797-824. doi: 10.1037/a0026187

Rossion, B., Delvenne, J.-F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guerit, J.-M. (1999). Spatio-temporal localization of the face inversion effect: an event-related potentials study. *Biological Psychology, 50,* 173-189. doi: 10.1016/S0301-0511(99)00013-7

Rossion, B., Gauthier, I., Tarr, M.J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport, 11*(1), 69-72.

Rossion, B., Joyce, C.A., Cottrell, G.W., & Tarr, M.J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage, 20(3),* 1609-1624. doi: 10.1016/j.neuroimage.2003.07.010

Rotge, J.Y., Lemogne, C., Hinfray, S., Huguet, P., Grynszpan, O., Tartour, E., et al. (2015). A meta-analysis of the anterior cingulate contribution to social pain. *Social, Cognitive & Affective Neuroscience, 10*(1), 19-27. doi: 10.1093/scan/nsu110

Rotshtein, P., Henson, R.N., Treves, A., Driver, J., & Dolan, R.J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience, 8*(1), 107-113. doi: 10.1038/nn1370

Rotter, J.B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35*(4), 651-665. doi: 10.1111/j.1467-6494.1967.tb01454.x

Rousseau, D.M., Sitkin, S.B., Burt, R.S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23*(3), 393-404. doi: 10.5465/AMR.1998.926617

Rousselet, G.A., Husk, J.S., Bennett, P.J., & Sekuler, A.B. (2005). Spatial scaling factors explain eccentricity effects on face ERPs. *Journal of Vision, 5*(10), 755-763. doi: 10.1167/5.10.1

Roy, A., Shepherd, S.V., & Platt, M.L. (2014). Reversible inactivation of pSTS suppresses social gaze following in the macaque (Macaca mulatta). *Social Cognitive & Affective Neuroscience, 9*(2), 209-217. doi: 10.1093/scan/nss123

Rudebeck, P.H., & Murray, E.A. (2014). The orbitofrontal oracle: Cortical mechanisms for the prediction and evaluation of specific behvioral outcomes. *Neuron, 84*(6), 1143-1156. doi: 10.1016/j.neuron.2014.10.049

Rudoy, J.D., & Paller, K.A. (2009). Who can you trust? Behavioral and neural differences between perceptual and memory-based influences. *Frontiers in Human Neuroscience, 3*(16), 1-6. doi: 10.3389/neuro.09.016.2009

Rule, N.O., Krendl, A.C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: behvioral and neural correlates. *Journal of Personality and Social Psychology, 104*(3), 409-426. doi: 10.1037/a0031050

Rule, N.O., Slepian, M.L., & Ambady, N. (2012). A memory advantage for untrustworthy faces. *Cognition, 125*(2), 207-218. doi: 10.1016/j.cognition.2012.06.017

Sabatinelli, D., Fortune, E.E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W.T., et al. (2011). Emotional perception: Meta-analyses of face and natural scene processing. *Neuroimage, 54*(3), 2524-2533. doi: 10.1016/j.neuroimage.2010.10.011

Sabatinelli, D., Keil, A., Frank, D.W., & Lang, P.J. (2013). Emotional perception: Correspondence of early and late event-related potentials with cortical and subcortical functional MRI. *Biological Psychology, 92*(3), 513-519. doi: 10.1016/j.biopsycho.2012.04.005

Sabatinelli, D., Lang, P.J., Keil, A., & Bradley, M.M. (2007). Emotional perception: correlation of functional MRI and event-related potentials. *Cerebral Cortex, 17*(5), 1085-1091. doi: 10.1093/cercor/bhl017

Sackeim, H.A., Greenberg, M.S., Weiman, A.L., Gur, R.C., Hungerbuhler, J.P., & Geschwind, N. (1982). Hemispheric asymmetry in the expression of positive and negative emotions. Neurologic Evidence. *Archives of Neurology, 39*(4), 210-218. doi: 10.1001/archneur.1982.00510160016003

Sadeh, B., Pitcher, D., Brandman, T., Eisen, A., Thaler, A., & Yovel, G. (2011). Stimulation of category-selective brain areas modulates ERP to their preferred categories. *Current Biology, 21*(22), 1894-1899. doi: 10.1016/j.cub.2011.09.030

Said, C.P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion, 9*(2), 260-264. doi: 10.1037/a0014681

Sánchez-Nàcher, N., Campos-Bueno, J.J., Sitges, C., & Montoya, P. (2011). Event-related brain responses as correlates of changes in predictive and affective values of conditioned stimuli. *Brain Research,1414,* 77-84. doi: 10.1016/j.brainres.2011.07.049

Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., & Cohen, J.D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755-1758. doi: 10.1126/science.1082976

Sato, W., Kochiyama, T., Uono, S., Matsuda, K., Usui, K., Inoue, Y., & Toichi, M. (2011). Rapid amygdala gamma oscillations in response to eye gaze. *PLoS One, 6*(11), 1-6. doi: 10.1371/journal.pone.0028188

Sato, W., Okada, T., & Toichi, M. (2007). Attentional shift by gaze is triggered without awareness. *Experimental Brain Research, 183*(1), 87-94. doi: 10.1007/s00221-007-1025-x

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *Neuroimage, 19*(4), 1835-1842. doi: 10.1016/S1053-8119(03)00230-1

Saxe, R., Xiao, D.K., Kovacs, G., Perrett, D.I., & Kanwisher, N. (2004). A region of the right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia, 42*(11), 1435-1446. doi: 10.1016/j.neuropsychologia.2004.04.015

Schachter, S., & Singer, J.E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review, 69,* 379-399. doi: 10.1037/h0046234

Scherer, K.S. (1999). Appraisal Theory. In T. Dalgleish and M.J. Power (Eds.), *Handbook of Cognition and Emotion* (pp.637-663). New York: Wiley.

Schilbach, L., Wilms, M., Eickhoff, S.B., Romanzetti, S., Tepest, R., Bente, G., Shah, N.J., Fink, G.R., & 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

Schotter, E.R., Berry, R.W., McKenzie, C.R., & Rayner, K. (2012). Gaze bias: Selective encoding and liking effects. *Visual Cognition, 18,* 1113-1132. doi: 10.1080/13506281003668900

Schug, J., Matsumoto, D., Horita, Y., Yamagishi, T., & Bonnet, K. (2010). Emotional expressivity as a signal of cooperation. *Evolution & Human Behaviour, 31,* 87-94. doi: 10.1016/j.evolhumbehav.2009.09.006

Schuller, A.M., & Rossion, B. (2001). Spatial attention triggered by eye gaze increases and speeds up early visual activity. *Neuroreport, 12*(11), 2381-2386.

Schuller, A.M., & Rossion, B. (2004). Perception of static eye gaze direction facilitates subsequent early visual processing. *Clinical Neurophysiology, 115*(5), 1161-1168. doi: 10.1016/j.clinph.2003.12.022

Schuller, A.M., & Rossion, B. (2005). Spatial attention triggered by eye gaze enhances and speeds up visual processing in upper and lower visual fields beyond early striate visual processing. *Clinical Neurophysiology, 116*(11), 2565-2576. doi: 10.1016/j.clinph.2005.07.021

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*, 1-27.

Schupp, H.T., Cuthbert, B.N., Bradley, M.M., Cacioppo, J.T., Ito, T., & Lang, P.J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology, 37*(2)*,* 257-261. doi: 10.1111/1469-8986.3720257

Schwarz, N., & Clore, G.L. (2003). Mood as information: 20 years later. *Psychological Inquiry, 14*(3-4), 296-303. doi: 10.1080/1047840X.2003.9682896

Schweinberger, S.R. (2011). Neurophysiological correlates of face recognition. In A.J. Calder, G. Rhodes, M.H. Johnson & J.V. Haxby (Eds.), *The Oxford handbook of face perception.* (pp, 345-366). New York: Oxford University Press.

Schweinberger, S.R., Kaufmann, J.M., Moratti, S., Keil, A., & Burton, A.M. (2007). Brain responses to repetitions of human and animal faces, inverted faces, and objects: an MEG study. *Brain Research, 1184,* 226-233. doi: 10.1016/j.brainres.2007.09.079

Schweinberger, S.R., Huddy, V. & Burton, A.M. (2004). N250r: a face selective brain response to stimulus repetitions. *Neuroreport, 15*(9), 1501-1505.

Schweinberger, S.R., Pickering, E.C., Jentzsch, I., Burton, A.M., & Kaufmann, J.M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research, 14*(3), 398-409. doi: 10.1016/S0926-6410(02)00142-8

Seidel, E.M., Silani, G., Metzler, H., Thaler, H., Lamm, C., Gur, R.C., et al. (2013). The impact of social exclusion vs. inclusion on subjective and hormonal reactions in females and males. *Psychoneuroendocrinology, 38*(12), 2925-2932. doi: 10.1016/j.psyneuen.2013.07.021

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

Seymour, B., O'Doherty, J.P., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., et al. (2004). Temporal difference models describe higher-order learning in humans. *Nature, 429*(6992), 664-667. doi: 10.1038/nature02581

Sharot, T., Velasquez, C.M., & Dolan, R.J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological Science, 21*(9), 1231-1235. doi: 10.1177/0956797610379235

Shepherd, S.V., Klein, J.T., Deaner, R.O., & Platt, M.L. (2009). Mirroring of attention by neurons in macaque parietal cortex. *Proceedings of the National Academy of Sciences USA, 106*(23), 9489-9494. doi: 10.1073/pnas.0900419106

Shultz, S., & Dunbar, R.I. M. (2007). The evolution of the social brain: anthropoid primates contrast with other vertebrates. *Proceedings of the Royal Society B: Biological Sciences, 274*(1624), 2429-2436. doi: 10.1098/rspb.2007.0693

Silberman, E.K., & Weingartner, H. (1986). Hemispheric lateralization of functions related to emotion. *Brain & Cognition, 5,* 322-353. doi: 10.1016/0278-2626(86)90035-7

Simmons, W.K., Ramjee, V., Beauchamp, M.S., McRae, K., Martin, A., & Barsalou, L.W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia, 45*(12), 2802-2810. doi: 10.1016/j.neuropsychologia.2007.05.002

Sims, T.B., Van Reekum, C.M., Johnstone, T., & Chakrabarti, B. (2012). How reward modulates mimicry: EMG evidence of greater facial mimicry of more rewarding happy faces. *Psychophysiology, 49*(7), 998-1004. doi: 10.1111/j.1469-8986.2012.01377.x

Slepian, M.L., Weisbuch, M., Adams, R.B., Jr., & Ambady, N. (2011). Gender moderates the relationship between emotion and perceived gaze. *Emotion, 11,* 1439-1444. doi: 10.1037/a0026163

Smith, A.P.R., Dolan, R. J., & Rugg, M.D. (2004). Event-Related Potential Correlates of the Retrieval of Emotional and Nonemotional Context. *Journal of Cognitive Neuroscience, 16*(5), 760-775. doi: 10.1162/089892904970816

Smith, E., Weinberg, A., Moran, T., & Hajcak, G. (2013). Electrocortical responses to NIMSTIM facial expressions of emotion. *International Journal of Psychophysiology, 88*(1), 17-25. doi: 10.1016/j.ijpsycho.2012.12.004

Sormaz, M., Andrews, T.J., & Young, A.W. (2013). Contrast negation and the importance of the eye region for holistic representations of facial identity. *Journal of Experimental Psychology: Human Perception & Performance, 39*(6), 1667-1677. doi: 10.1037/a0032449

Speckmann, E-J., Elger, C.E., & Gorji, A. (2011). Neurophysiological basis of EEG and DC potentials. In D. L. Schomer & F. Lopes da Silva (Eds.), *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields (Sixth Edition),* (pp. 17-32).

Spielberger, C.D. (1983). *Manual for the State-Trait Anxiety Inventory (STAI).* PaloAlto, CA: Consulting Psychologists Press.

Stirrat, M., & Perrett, D.I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science, 21*(3), 349-354. doi: 10.1177/0956797610362647

Stouten, J., & De Cremer, D. (2009). “Seeing is believing”: The effects of facial expressions of emotion and verbal communication in social dilemmas. *Journal of Behavioral Decision Making, 23*(3), 271-287. doi: 10.1002/bdm.659

Striano, T., & Reid, V.M. (2006). Social cognition in the first year. *Trends in Cognitive Sciences, 10*(10), 471-476. doi: 10.1016/j.tics.2006.08.006

Süßenbach, F., & Schönbrodt, F. (2014). Not afraid to trust you: Trustworthiness moderates gaze cueing but not in highly anxious participants. *Journal of Cognitive Psychology, 26*(6), 670-678. doi: 10.1080/20445911.2014.945457

Suzuki, A., Honma, Y., & Suga, S. (2013). Indelible distrust: memory bias toward cheaters revealed as high persistence against extinction. *Journal of Experimental Psychology: Learning, Memory & Cognition, 39*(6), 1901-1913. doi: 10.1037/a0033335

Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science, 22*(6), 430-436. doi: 10.1177/0963721413498893

Talmi, D., & McGarry, L.M. (2012). Accounting for immediate emotional memory enhancement. *Journal of Memory and Language, 66*(1), 93-108. doi: 10.1016/j.jml.2011.07.009

Tamietto, M., Castelli, L., Vighetti, S., Perozzo, P., Geminiani, G., Weiskrantz, L., & d Gelder, B. (2009). Unseen facial and bodily expressions trigger fast emotional reactions. *Proceedings of the National Academy of Sciences USA, 106*(42), 17661-17666. doi: 10.1073/pnas.0908994106

Tanaka, J.W., Curran, T., Porterfield, A.L., & Collins, D. (2006). Activation of pre-existing and acquired face representations: the N250 event-related potential as an index of face familiarity. *Journal of Cognitive Neuroscience, 18*(9), 1488-1497. doi: 10.1162/jocn.2006.18.9.1488

Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 46*(2), 225-245. doi: 10.1080/14640749308401045

Tassinary, L.G., Cacioppo, J.T., & Vanman, E.J. (2007). The skeletomotor system: Surface electromyography. In J.T. Cacioppo, L.G. Tassinary, & G.G. Bernston, *Handbook of Psychophysiology,* (pp. 267-299). New York: Cambridge University Press.

Teufel, C., Alexis, D.M., Clayton, N.S., & Davis, G. (2010). Mental-state attribution drives rapid, reflexive gaze following. *Attention, Perception & Psychophysics, 72*(3), 695-705. doi: 10.3758/APP.72.3.695

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review, 9*(2), 314-318. doi: 10.3758/BF03196287

Tipples, J. (2008). Orienting to counterpredictive gaze and arrows cues. *Perception & Psychophysics, 70*(1), 77-87. doi: 10.3758/PP.70.1.77

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

Todd, R.M., Talmi, D., Schmitz, T.W., Susskind, J., & Anderson, A.K. (2012). Psychophysical and neural evidence for emotion-enhanced perceptual vividness. *Journal of Neuroscience, 32*(33), 11201-11212. doi: 10.1523/JNEUROSCI.0155-12.2012

Todorov, A., Baron, S., & Oosterhof, N.N. (2008). Evaluating face trustworthiness: A model based approach. *Social, Cognitive & Affective Neuroscience, 3,* 119-127. doi: 10.1093/scan/nsn009

Todrank, J., Byrnes, D., Wrzesniewski, A., & Rozin, P. (1995). Odors can change preferences for people in photographs: A cross-modal evaluative conditioning study with olfactory USs and visual CSs. *Learning & Motivation, 26*(2), 116-140. doi: 10.1016/0023-9690(95)90001-2

Tomarken, A.J., Davidson, R.J., Wheeler, R.E., & Doss, R.C. (1992). Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *Journal of Personality and Social Psychology, 62*(4), 676-687. doi: 10.1037/00223514.62.4.676

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28,* 675-735. doi: 10.1017/S0140525X05000129

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

Tranel, D., Damasio, H., & Damasio, A.R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience, 7*(4), 425-432. doi: 10.1162/jocn.1995.7.4.425

Tranel, D., Damasio, H., & Damasio, A.R. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia, 35*(10), 1319-1327. doi: 10.1016/S0028-3932(97)00085-7

Treinen, E., Corneille, O., & Luypaert, G. (2012). L-eye to me: the combined role of Need for Cognition and facial trustworthiness in mimetic desires. *Cognition, 122*(2), 247-251. doi: 10.1016/j.cognition.2011.10.006

Trimble, M. (2010). Sodium Amytal Testing of the Laterality of Emotion. *Clinical EEG and Neuroscience, 41*(4), 211-213. doi: 10.1177/155005941004100408

Tsao, D.Y., Moeller, S., & Freiwald, W. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of science USA, 105*(49)*,* 19514-19519. doi: 10.1073/pnas.0809662105

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Parformance, 24*(3), 830-846. doi: 10.1037/0096-1523.24.3.830

Tummeltshammer, K.S., Wu, R., Sobel, D.M., & Kirkham, N.Z. (2014). Infants track the reliability of potential informants. *Psychological Science, 25*(9), 1730-1738. doi: 10.1177/0956797614540178

Tversky, A., & Kahneman, D. (1984). Choices, Values, and Frames. *American Psychologist, 39*(4), 341-350. doi: 10.1037/0003-066X.39.4.341

Ulloa, J.L., Marchetti, C., Taffou, M., & George, N. (2015). Only your eyes tell me what you like: Exploring the liking effect induced by other's gaze. *Cognition & Emotion, 29*(3), 460-470. doi: 10.1080/02699931.2014.919899

Uskul, A., & Over, H. (2014). Responses to social exclusion in cultural context: Evidence from farming and herding communities. *Journal of Personality and Social Psychology, 106*(5), 752-771. doi: 10.1037/a0035810

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology A, 43*(2), 161-204. doi: 10.1080/14640749108400966

van Boxtel, A. (2001). Optimal signal bandwidth for the recording of surface EMG activity of facial, jaw, oral, and neck muscles. *Psychophysiology, 38,* 22-34. doi: 10.1111/1469-8986.3810022

van Honk, J., Eisenegger, C., Terburg, D., Stein, D.J., & Morgan, B. (2013). Generous economic investments after basolateral amygdala damage. *Proceedings of the National Academy of Sceinces 110*(7), 2507-2510. doi: 10.1073/pnas.1217316110

van't Wout, M., Kahn, R.S., Sanfey, A.G., & Aleman, A. (2006). Affective state and decision-making in the ultimatum game. *Experimental Brain Research, 169*(4), 564-568. doi: 10.1007/s00221-006-0346-5

Vogt, B.A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nature Reviews Neuroscience, 6*(7), 533-544. doi: 10.1038/nrn1704

Vuilleumier, P., & Driver, J. (2007). Modulation of visual processing by attention and emotion: windows on causal interactions between human brain regions. *Philosophical Transactions of the Royal Society B, 362,* 837-855. doi: 10.1016/S0079-6123(06)56004-2

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia, 45*(1), 174-194. doi: 10.1016/j.neuropsychologia.2006.06.003

Watson, K.K., & Platt, M.L. (2012). Social signals in primate orbitofrontal cortex. *Current Biology, 22*(23), 2268-2273. doi:10.1016/j.cub.2012.10.016

Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: the time-course of neural activity elicited by specific picture content. *Emotion, 10*(6), 767-782. doi: 10.1037/a0020242

Weiskrantz, L. (1956). Behavioral changes associated with ablation of the amygdaloid complex in monkeys. *Journal of Comparative and Physiological Psychology, 49*(4), 381-391.

Wheeler, R.E., Davidson, R.J., & Tomarken, A.J. (1993). Frontal brain asymmetry and emotional reactivity: a biological substrate of affective style. *Psychophysiology, 30*(1), 82-89. doi: 10.1111/j.1469-8986.1993.tb03207.x

Wiese, H., Altmann, C.S., & Schweiberger, S. (2014). Effects of attractiveness on face memory separated from distinctiveness: Evidence from event-related brain potentials. *Neuropsychologia, 56,* 26-36. doi: 10.1016/j.neuropsychologia.2013.12.023

Williams, K.D. (2007). Ostracism. *Annual Review of Psychology, 58,* 425-452. doi: 10.1146/annurev.psych.58.110405.085641

Williams, K.D., Cheung, C.K., & Choi, W. (2000). Cyberostracism: effects of being ignored over the internet. *Journal of Personality & Social Psychology, 79*(5), 748-762. doi: 10.1037/0022-3514.79.5.748

Willis, J., & Todorov, A. (2006). First impressions: making up your mind after 100-ms exposure to a face. *Psychological Science, 17*(7), 592-598. doi: 10.1111/j.1467-9280.2006.01750.x

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G.R., & Vogeley, K. (2010). It's in your eyes – using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social, Cognitive and Affective Neuroscience, 5*(1), 98-107. doi: 10.1013/scan/nsq024

Wincenciak, J., Dzhelyova, M., Perrett, D.I., & Barraclough, N.E. (2013). Adaptation to facial trustworthiness is different in female and male observers. *Vision Research, 87,* 30-34. doi: 10.1016/j.visres.2013.05.007

Winkielman, P., & Cacioppo, J.T. (2001). Mind at ease puts a smile on the face: Physiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology, 81*(6), 989-1000. doi: 10.1037/0022-3514.81.6.989

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science, 17*(9), 799-806. doi: 10.1111/j.1467-9280.2006.01785.x

Winston, J.S., Strange, B.A., O'Doherty, J., & Dolan, R.J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience, 5*(3), 277-283. doi: 10.1038/nn816

Wood, S. & Kisley, M.A. (2006). The negativity bias is eliminated in older adults: age-related reduction in event-related brain potentials associated with evaluative categorization. *Psychology and Aging, 21*(4), 815-820. doi: 10.1037/0882-7974.21.4.815

Wubben, M.J.J., De Cremer, D., & van Dijk, E. (2011). The communication of anger and disappointment helps to establish cooperation through indirect reciprocity. *Journal Economic Psychology, 32*(3), 489-501. doi: 10.1016/j.joep.2011.03.016

Xu, S., Zhang, S., & Geng, H. (2011). Gaze-induced joint attention persists under high perceptual load and does not depend on awareness. *Vision Research, 51*(18)*,* 2048-2056. doi: 10.1016/j.visres.2011.07.023

Yang, D., Qi, S., Ding, C., & Song, Y. (2011). An ERP study on the time course of facial trustworthiness appraisal. *Neuroscience Letters, 496*(3), 147-151. doi: 10.1016/j.neulet.2011.03.066

Yin, R.K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141-145. doi: 10.1037/h0027474

Young, A.W., Hellawell, D., & Hay, D.C. (1987). Configural information in face perception. *Perception, 16*(6), 747-759. doi: 10.1068/p160747

Young, A.W., Perrett, D.I., Calder, A.J., Sprengelmeyer, R., & Ekman, P. (2002). *Facial Expressions of Emotion: Stimuli and Test (FEEST).* Bury St. Edmunds: Thames Valley test Company.

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2), 1-27. doi: 10.1037/h0025848

Zheng, X., Mondloch, C.J., & Segalowitz, S.J. (2012). The timing of individual face recognition in the brain. *Neuropsychologia, 50*(7), 1451-1461. doi: 10.1016/j.neuropsychologia.2012.02.030

Zimmerman, F.G., & Eimer, M. (2013). Face learning and the emergence of view-independent face recognition: an event-related brain potential study. *Neuropsychologia, 51*(7), 1320-1329. doi: 10.1016/j.neuropsychologia.2013.03.028