The integration of facial features over space and time

Dissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

der Mathematisch-Naturwissenschaftlichen Fakultät
und
der Medizinischen Fakultät
der Eberhard-Karls-Universität Tübingen

vorgelegt
von

Maren Reinl
aus Rostock, Deutschland

Mai 2016

Tag der mündlichen Prüfung: 17.06.2016

Dekan der Math.-Nat. Fakultät: Prof. Dr. W. Rosenstiel
Dekan der Medizinischen Fakultät: Prof. Dr. I. B. Autenrieth

1. Berichterstatter: Dr. Andreas Bartels

2. Berichterstatter: Prof. Dr. Dirk Wildgruber

Prüfungskommission: Prof. Dr. Martin Giese

Dr. Andreas Bartels

Prof. Dr. Dirk Wildgruber

Prof. Dr. Hubert Preißl

Erklärung / Declaration:

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

„………………………………………………………………………………………..................
..............................……...…“

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

I hereby declare that I have produced the work entitled "……", submitted for the award of a doctorate, on my own (without external help), have used only the sources and aids indicated and have marked passages included from other works, whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.

Tübingen, den ....

Datum / Date Unterschrift /Signature

Augen in der Großstadt

Kurt Tucholsky

Wenn du zur Arbeit gehst
am frühen Morgen,
wenn du am Bahnhof stehst
mit deinen Sorgen:
da zeigt die Stadt
dir asphaltglatt
im Menschentrichter
Millionen Gesichter:
Zwei fremde Augen, ein kurzer Blick,
die Braue, Pupillen, die Lider -
Was war das? vielleicht dein Lebensglück…
vorbei, verweht, nie wieder.

Du gehst dein Leben lang
auf tausend Straßen;
du siehst auf deinem Gang, die
dich vergaßen.
Ein Auge winkt,
die Seele klingt;
du hast's gefunden,
nur für Sekunden…
Zwei fremde Augen, ein kurzer Blick,
die Braue, Pupillen, die Lider -
Was war das? Kein Mensch dreht die Zeit zurück…
Vorbei, verweht, nie wieder.

Du mußt auf deinem Gang
durch Städte wandern;
siehst einen Pulsschlag lang
den fremden Andern.
Es kann ein Feind sein,
es kann ein Freund sein,
es kann im Kampfe dein
Genosse sein.
Er sieht hinüber
und zieht vorüber …
Zwei fremde Augen, ein kurzer Blick,
die Braue, Pupillen, die Lider -
Was war das?
Von der großen Menschheit ein Stück!
Vorbei, verweht, nie wieder.

— A chapter of my life. Many thanks to every friend I met on that journey. —

Abstract

Faces are unique social stimuli that can be recognized in an instant. We can pick up information about gender, ethnicity, feelings, attentional focus or even attributes like attractiveness or trustworthiness remarkably quickly. How we achieve this has been subject to psychology, cognitive science and neuroscience since decades but we still don´t know the full picture. The key theme of this thesis concerns the integration, or binding, of facial features over space and over time. We investigated both behavioral measures in healthy people and a group of people with Autism Spectrum Disorder (ASD), and the neuronal mechanisms in core face processing regions of the human brain.

The first part of this thesis investigates the contribution of face responsive brain areas to whole face and part-based neural representation of facial expressions. This aspect has hardly been considered in the past, as most studies focused on the representation of identity instead. During a fmri-experiment, we presented whole faces and facial parts of happy and fearful expressions. We extracted the similarity of activity patterns in core network of face processing - occipital face area (OFA), fusiform face area (FFA) and superior temporal sulcus (STS) - across and within emotions between whole faces and facial parts. Previous studies based on identitity recognition have found holistic and part-based representations in the FFA while the OFA seems to mainly represent part-based information. The STS has hardly been considered in those studies, as it is thought to be preferentially involved to expression coding. We find both part-based representation of facial expressions and an emotion-indepented preference of whole faces in the FFA, in line with the previous findings for identity recognition. For STS, we detect emotion-dependent representations of faces and facial parts, supporting its major role in expression processing. The OFA, in contrast, shows similar representation of the eyes- and mouth-regions of both expressions without any further specific effects, adding evidence to its role as an entry-point of facial information into the core network of face processing.

The second part of the thesis explores the temporal information embodied in dynamic facial expressions. Using expressions of increasing and decreasing intensity that were presented in the natural or reversed frame order, we manipulate the temporal information of expression unfolding in a well controlled 2x2 Design (factor “emotion direction” and factor “timeline”). This approach allowed us to nicely control for low-level aspects. In three consecutive studies, we explore first the underlying brain activation elicited by our stimulus manipulation in healthy subjects. Second, we examine the perceptual effects caused by emotion-direction and timeline-reversal in healthy subjects, and third in autistic participants and matched controls. Our results indicate a sensitivity of all areas of the neural core network of face processing to both, emotion-direction and timeline. Behaviorally, we found that both factors affected jugdements of different stimulus properties like emotion intensity or how well emotions are performed, even if subjects were not informed of the timeline manipulation. Interestingly, autistic subjects did not differ from the control group regarding the effects caused by timeline reversal in their perceptual evaluation of the stimuli.

In sum, our studies shed light onto two key aspects of facial processing and perception - holistic or part-based processing and facial dynamics - that have not been addressed before in the way done here.

Content

1 Synopsis

Every single day of our live, we come across various different faces - when we greet our family in the morning, when we take the bus to work, when we walk along the streets, meet up with friends for a chat and even when we watch a movie in the evening. We easily recognize people that we know and we can guess the state of mind from people we have never seen before. We might even fall in love with a pair of eyes that meets our glance. From a scientific point of view, the ability to extract such large amount of information from faces so effortlessly is fascinating. How we actually achieve that, has driven research in psychology, cognitive science and neuroscience since decades. Several striking phenomena, ranging from an early face preference in newborns (Carey, 1981; Ellis, 1990) along impressive effects like the findings for face inversion (Yin 1969) or the composite face effect (Galton, 1879; Young et al., 1987) to the isolated disturbances of face perception in prosopagnosia (Bodamer, 1947; see Ellis and Florence, 1990 for a partial translation), have demonstrated the specialty of faces compared to other visual stimuli. Although research has come a long way to understand a lot about how we perceive and process faces, there are still many open questions. This thesis addresses two topical questions in the current face-research concerning the spatial and temporal integration of facial features. First, we investigated the representation of facial expressions in brain areas dedicated to face processing, an issue that was long neglected while focusing on the representation of facial identity instead. Second, we present some of the first studies investigating neural and behavioral responses to facial dynamics under precisely controlled conditions.

1.1 Representation of faces: parts-based versus holistic perception

> *A man shall see faces, that if you examine them part by part, you shall find never a good; and yet altogether do well. (Francis Bacon (1561-1626): Of Beauty)*

Already in 1879, Galton (1879) combined pieces of different faces into one percept and found those composite faces to appear real and even better looking than their components. Today, the effect that features from different faces together appear as a new face is called 'holistic representation'. However, by now there are many different definitions and experimental operations how this holistic percept is formed (Richler et al., 2012). Often, holistic processing is described as the perception of a face as one 'Gestalt' without much decomposition into its parts (Young et al., 1987; Tanaka and Farah, 1993; Farah et al., 1998; Bartlett et. al., 2003; Rossion, 2008), while others rather consider configural properties, which describe the T-shaped arrangement (first-order relations) and the metric distances between eyes, nose, mouth and other facial features (second-order relations) as underlying source of holistic information (Diamond and Carey, 1986; Rhodes et al., 1993; Searcy and Bartlett, 1996; Leder and Bruce, 2000). Again, some propose that holistic representations emerge from the interaction of featural (refers to size, colour, form et cetera of facial parts) and configural information (Amishav and Kimchi, 2010; Kimchi and Amishav, 2010). Fourthly, there is the idea that the holistic effect is based on the high experience with faces that automates the perceptual process (Wong and Gauthier, 2010; Richler et al., 2011a; Richler et al., 2011b). Many studies investigated the impact of configural and featural face information and their integration into a holistic percept, but the results seem to depend at least partly on the experimental design and experimental operationalization (Maurer et al., 2002).

One of the first effects that indicates the specialty of faces is the so-called inversion effect (Yin, 1969). Recognizing faces is stronger impaired than recognizing other objects when stimuli are turned upside down. The composite face effect (Young et al., 1987) further adds evidence that faces are processed as a whole with single features affecting the perception of other features. Another famous task, the part-whole-effect, introduced by Tanaka and Farah (1993), shows that the memory for facial parts is better when they are learned in the context of a whole face than when presented in isolation. This is not the case when testing other objects like houses. This suggests that the memory for facial features is integrated into a whole face representation. If the spatial distances between single features are varied in the recognition test, performance decreases, which further indicates the importance of second-order configural information for face recognition. These paradigms support the idea that faces are presented in a holistic fashion and are therefore special in comparison to other objects.

1.2 The neuronal representation of faces

The currently most influential model of the neural representation of faces was proposed by Haxby et al., (2000). They propose a distributed system for face perception that distinguishes between a 'Core network' in the visual extrastriate areas, executing primarily the visual analysis of faces, as well as an 'Extended system' performing further attention-, speech-, emotion- or biographical-related analysis. The core system consists of three areas: The occipital face area (OFA) in the inferior occipital gyrus, the fusiform face area (FFA) in the lateral fusiform gyrus and the superior temporal sulcus (STS). The OFA is thought to represent an entry point of facial information into the core system (Pitcher et al., 2007; Pitcher et al., 2011b), providing a first visual analysis of the facial information. However, this proposition is open to debate by other findings (Kadosh et al., 2011). From there, the route departs into two branches. A distinction is made between invariant and changeable aspects of faces. Invariant aspects refer to non-changeable properties like the identity of a face, which are supposed to be processed mainly in the FFA (Furl et al., 2011; Hoffman and Haxby, 2000; Nestor et al., 2011). In contrast, changeable aspects like facial expressions, eye-gaze or speech-related mouth movements are described to be processed in the STS (Puce et al., 1998; Said et al., 2010; Hoffman and Haxby, 2000; Campbell et al., 2001). The distinction of changeable and non-changeable aspects of faces has already been implicated by the model of Bruce and Young (1986). Several reports of selective impairment of identity recognition after brain injuries go in line with this model. However, it is unclear what exactly causes those injury-related impairments. Alongside distinctive processing routes of identity and expression, a selective impairment of facial memory function could also be the underlying cause (see Calder et al., 2000 for a detailed discussion). The strict distinction within the core network proposed by Haxby et al., (2000) has further been challenged by other findings. The FFA has repeatedly been shown to respond stronger to emotional compared to neutral faces, even when controlling for attentional effects (Dolan et al., 1996; Gerber et al., 2008; Ishai et al., 2004; Pessoa et al., 2002; Surguladze et al., 2003; Vuilleumier and Pourtois, 2007; Vuilleumier et al., 2001; Winston et al., 2003). In contrast, some studies failed to report STS activation towards facial expressions (George et al., 1993; Sergent et al., 2007); while others clearly show its involvement in decoding expressional information (Said et al., 2010). Furthermore, detected dependencies of expression- and speech-recognition on identity as well as the enhancing effects of motion on identity recognition have suggested a stronger interaction of both processing routes (Schweinberger and Soukup, 1998; Schweinberger et al., 1999). It is quite likely, that the processing of identity and expression does occur in separate mechanisms but that they are not fully independent from each other but reflect a 'bias rather than a categorical dissociation' (Calder and Young, 2005).

1.3 Holistic versus part-based processing in the brain

The model of Haxby et al., (2000) does not make assumptions about holistic and part-based representations in the areas of the core network. By now, there is a bunch of studies targeting this question. The FFA is described as representing whole faces and configural as well as featural information while the OFA is supposed to mainly represent featural and, only to a smaller degree, holistic information. Comparing houses and faces in a task where subjects either had to detect changes on featural or global level, Rossion (Rossion et al., 2000) showed that left FFA represented featural face information while right FFA responded stronger to whole faces. In contrast, Yovel and Kanwisher (2004) found no hemispheric differences of featural and configural information in the FFA. Rotshtein et al., (2007) showed sensitivity of the FFA to the converged information of featural and configural information. Liu et al., (2010) and Golarai et al., (2015) additionally report sensitivity for real face parts compared to schematic parts in FFA while the OFA only responded to the naturalness of the stimuli, independent of the arrangement of the facial features. Release from adaptation to upright composite face images compared to misaligned or inverted faces was found for FFA (Schiltz and Rossion, 2006; Schiltz et al., 2010) and to a lesser extent for OFA (Schiltz and Rossion, 2006), indicating holistic processing in both areas. In 2015, Soria Bauser and Suchan (2015) compared the responses of FFA and OFA to whole faces as well as scrambled faces, and found stronger activation to whole faces compared to scrambled faces for FFA and left OFA, but the opposite pattern for right OFA. As their scrambling methods did not scramble face parts within a face, but instead cut the whole picture into pieces, they used a different manipulation than previous studies. They also disturbed the impression of a coherent picture and not only the first-order-relationship of the facial parts.

However, the above-mentioned studies are based on identity recognition, a reason why the STS was often not considered (Rhodes et al., 2009). Given the proposed neuronal distinction as well as differential effects found for holistic and part-based representations of facial identity and expression, it is unclear whether the findings also hold true when studying the representation of expressional faces. Behavioral results indicate a stronger part-based representation of facial expressions (McKelvie, 1995; Prkachin, 2003; Calvo and Nummenmaa, 2008; Derntl et al., 2009; Narme et al., 2011; Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011; Gosselin and Schyns, 2001; Blais et al., 2012) compared to the rather holistic representation face identity. To extend the research on the neuronal representation of facial expressions might help to clarify the distinction of expression and identity processing in the brain. Surprisingly, only recently researchers started to address this question. However, first results indicate a representation of facial expression similar to the way identity is represented in the brain (Calvo and Beltrán, 2014). In two studies, after all, the STS was included in the analysis and seems to represent facial expressions mainly in a part-based fashion (Said et al., 2010; Flack et al., 2015).

1.4 Dynamic face stimuli

In our daily life, we usually encounter moving instead of static faces. We see other people talk, direct their attention by the shift of eye-gaze or moving their heads, we see emotional expressions arising and dropping off and we effortlessly process all this information in an instant. Still, in research the majority of experiments were done using face photographs for decades. The use of dynamic faces as more ecologically valid increased only during the past years. Such dynamic face stimuli have introduced a totally new dimension into face processing research: time and temporal information. This information has totally been neglected before.

While at the beginning, most studies compared dynamic and static face stimuli against each other, the recent technical developments now have widened the options of creating well designed control stimuli which is a necessary precondition to systematically study moving faces. Dynamic face stimuli have been shown to improve the recognition of emotional expressions (Ambadar et al., 2005; Wehrle et al., 2000) and can even help to discriminate gender and identities of faces (O'Toole et al., 2002; Thornton and Kourtzi, 2002). Hill and Johnston (2001) used a very interesting approach to test the impact of facial motion. They used movement patterns of different people to animate a standard head. Subjects were able to detect the correct identities just from the facial dynamics. Additionally, the core network of face processing increases its activity when processing dynamic compared to static faces (Fox et al., 2009; Kilts et al., 2003; Labar et al., 2003; Sato et al., 2004; Schultz and Pilz, 2009; Trautmann et al., 2009). The biggest rise in activity has been found for the posterior STS (Pitcher et al., 2011; Schultz et al., 2013). This finding is not surprising given its involvement in biological motion processing (Bonda et al., 1996; Peelen et al., 2006; Blake and Shiffrar, 2007; Decety, 1999).

However, we still lack a full understanding of what information is actually provided in the motion that reaches beyond information that we can extract already from still pictures. Of course, it could be the higher amount of static information presented in the movies, as they are in principle sequences of static pictures presented with a certain frame rate. However, also the fluidity and meaningfulness of the dynamics have been shown to be necessary for the advantage of dynamic face stimuli. Scrambled movies neither enhance expression recognition nor increase the activity in the core regions of face processing (Schultz et al., 2013; Furl et al., 2010). In fact, scrambling was often used as a more proper control condition than static faces. Another way to generate highly controlled stimuli is the use of facial morphs (Sato et al., 2010). The disadvantage of morphs, however, is that they usually contain a linear change of all facial features at the same time. Unfortunately, this might not resemble the dynamic changes occurring in the real life.

The Facial Action Coding System (FACS) has been developed to describe changes in facial features for different emotions (Ekman et. al., 2002). It defines so-called action units (AU) which represent different muscles and features in the face. Using these AU, prototypical patterns of muscle activations of facial expressions have been described, however, only for the peak expression in static stimuli (Ekman et. al., 2002). It is very likely that in the natural evolvement of facial expressions, the AUs show different time courses for different emotions. Curio et. al., (2006) developed a morph algorithm based on facial motion previously tracked from real participants. This technique shows different temporal patterns for different AUs. They compared this type of morphs to a global morph, animating the whole face at the same time. They showed sensitivity of subjects to the small variations in movement timing. Participants rated morphed stimuli as more natural when the real spatio-temporal information was taken into account as when the whole face was animated at once. This underlines the importance of considering natural facial dynamics in research.

1.5 Disturbances of face perception

Several clinical conditions involve disturbances in face processing which can lead to dramatic consequences in social interaction. The most prominent disorder is prosopagnosia, also called face blindness, where patients are impaired to recognize mainly facial identity (Bodamer, 1974; see Ellis and Florence, 1990 for a partial translation). Less obvious disturbances can be found in psychiatric conditions like autism. Autism among others is characterized by impairments in social interactions and communication and the diverging clinical manifestations of autistic traits are summarized as Autism Spectrum Disorder (ASD).

One reason of the deficits in social contact found in ASD seems to be the diminished interest in faces. One of the early symptoms is the lack of attention towards eye-contact or the reduced following of gaze shifts from other people. However, there is a high variability of performance on facial tasks in subjects with ASD. Many studies on identifying or matching facial expressions have found reduced performance in ASD participants (Ashwin et al., 2007; Critchley et al., 2000; Malisza et al., 2011); while others did not find any differences in comparison to control groups (Bird et al., 2006; Dapretto et al., 2006; Kleinhans et al., 2010; Wicker et al., 2008; Castelli, 2005). Corbett et al., (2009) found impairments for matching identity when expressional faces were used, but no impairments for expression matching itself. However, sometimes longer reaction times have been reported for subjects with ASD (Piggot et al., 2004; Corbett et al., 2009; Kleinhans et al., 2009), indicating difficulties in facial tasks despite similar recognition accuracy. When matching facial expressions with voices, Hobson repeatedly found impairments for ASD participants (Hobson, 1986a; Hobson, 1986b; Hobson et al., 1988). Matching expressional faces and words, however, leads to none or only small differences (Piggot et al., 2004; Wang et al., 2004). In general, findings seem to largely depend on the task of the experiment (see Maurer et al., 2002 for a detailed review) but there are undoubted impairments in ASD when faces need to be processed and evaluated.

1.5.1 Neural underpinnings of disturbed face processing

Very often, activity pattern and functional connectivity of different brain areas involved in face perception and social cognition are altered in ASD (Malisza et al., 2011). The main finding is a hypoactivation for FFA (Critchley et al., 2000; Deeley et al., 2007; Wang et al., 2004; Corbett et al., 2009; Humphreys et al., 2008; Pierce et al., 2001; Domes et al., 2013) which often goes in line with reduced amygdala activity (Domes et al., 2013; Hadjikhani et al., 2007; Pierce et al., 2001). Furthermore, the OFA (Humphreys et al., 2008; Pierce et al., 2001; Domes et al., 2013) and the STS (Ashwin et al., 2007; Hubl et al., 2003; Wicker et al., 2008; Hadjikhani et al., 2007; Humphreys et al., 2008; Pierce et al., 2001) have been found to be less activated in subjects with ASD. In contrast, some studies do not report such differences between subjects with and without ASD (Hadjikhani et al., 2004; Hadjikhani et al., 2007). Pierce and Redcay (2008) showed normal FFA activation for familiar but reduced FFA activation for unfamiliar faces. Corbett et al., 2009 found reduced FFA activity for identity matching but reduced amygdala activation for emotion matching.

In the past years, the underlying neural basis of impaired social cognition and face processing has increasingly been attributed to abnormal connectivity patterns in the brain rather than the malfunctioning of single areas (Horwitz et al., 1988; Belmonte et al., 2004; Wickelgren, 2005). Altered connectivity of face- and social cognition-related areas have been found using DTI (Barnea-Goraly et al., 2004), structural modelling (Wicker et al., 2008) or correlation analysis (Koshino et al., 2008). Also, cognitive and language impairments in ASD have been related to such disturbances (see Maximo et al., 2014 for review). Repeatedly, the amygdala is described within these dysfunctional neural networks (Monk et al., 2010; Murphy et al., 2012; Schultz, 2005). This is often reflected in atypical co-activation of FFA and amygdala. Reduced amygdala activation has further been reported to be correlated with reduced exploration of the eye-region in a face. However, it is unclear if this diminished amygdala activity is caused by an active avoidance of the eye contact, or if the amygdala rather triggers orientation towards the eyes and therefore reduced amygdala activity results in less eye contact. There is evidence for both explanations. Dalton et al., (2005) found amygdala activity to be positively correlated with eye contact. They interpret this rise in amygdala activity as a consequence of the aversiveness of eye contact in ASD, supporting the active-avoidance-hypothesis. This explanation might also account for findings of higher amygdala activity in some studies of face processing with ASD participants (Ashwin et al., 2007; Weng et al., 2011; Monk et al., 2010). Support for the second hypothesis of a lack of attention following hypoactivation of the amygdala arises from patients with amygdala lesions. They have been found to show less reflexive orientation towards the eyes (Adolphs et al., 2005; Spezio et al., 2007) while in healthy subjects an increase in amygdala activity seems to be connected to enhanced orientation toward the eyes (Gamer and Büchel, 2009; Gamer et al., 2013).

1.5.2 The use of dynamic faces to study disturbed face processing

Unfortunately, the use of dynamic face stimuli has not revealed much new information about the face processing deficits in ASD, as inconsistent results persist (Loveland et al., 1997; Gepner et al., 2001; Lindner and Rosén, 2006; Tardif et al., 2007; Bal et al., 2010; Grossman and Tager-Flusberg, 2008). One reason might be potential motion processing impairments, especially when it comes to biological and human motion. In recent years, a growing number of studies have investigated the perception of biological motion in ASD, often using point-light displays (PLD) of human bodies. Results have been again inconsistent across studies, which may be also due to the various design of each study. Testing the detection of human motion by presenting intact and scrambled Point-light-walker (PLW) (Freitag et al., 2008; Murphy et al., 2009; Annaz et al., 2012), similar detection rates for ASD and typical developed (TD) participants but slower reaction times for ASD had been reported. Freitag et al., (2008) however found also slower performance on scrambled PLW, leaving it unclear whether this effect is due to a general visuo-motor impairment in ASD. The first study using PLDs of human movement was conducted by Moore et. al., (1997). ASD children showed similar performance than TD to report the action of the PLW but performed worse on emotion recognition. They concluded that ASD show normal sensitivity to human movement but are impaired when emotional expressional are involved. Similar results have also been reported in later studies (Parron et al., 2008; Hubert et al., 2007) but also contrasting results have been found (Atkinson, 2009). However, it is unclear whether the detected differences can be related to language abilities as participants had to verbally name the demonstrated expression. Blake et. al., (2003) found decreased performance on PLW-task to correlate significantly with autism severity, a finding later confirmed by Kaiser and Pelphrey (2012) and Wright et al., (2014). Other studies found additional correlations of performance on biological motion tasks with non-verbal age (Rutherford and Troje, 2012) and IQ (Jones et al., 2011; Rutherford and Troje, 2012). Rutherford and Troje, (2012) suggest that the development of alternative strategies might increase performance in higher-IQ ASD, as differential brain responses are still reported despite similar performance in PLW-tasks. Additionally, Kaiser and Pelphrey (2012) and Wright et al., (2014) found autistic traits to be correlated with the detection of human but not object motion. When comparing attention and performance of ASD and TD on object motion compared to biological motion, ASD do not prefer biological motion and show less impairments for object motion (Annaz et al., 2012). This is comparable to similar performance on object and face tasks and a diminished interest in faces.

1.6 Thesis Overview

The goal of this thesis was to target two important, timely, and much under-studied questions in the areas of face research.

First, it aims to deepen the understanding of how faces and facial parts are represented in the brain and integrated into a single percept. This question was approached with a fMRI study using a pattern-based multivariate analysis technique presented in Chapter two.

Second, it attempts to extend the scarce knowledge of how facial dynamics are perceived and processed in the brain. Three successive studies were conducted to investigate the impact of facial dynamics and temporal information for face perception. Chapter three describes a fMRI study and Chapter four presents the corresponding behavioral data. The final Chapter, Chapter five then extends the previous behavioral findings by comparing the performance of participants with and without ASD.

The two questions, concerning holistic and dynamic face processing are highly related: in both, we attempt to find neuronal underpinnings for integrative processes. In the first, the question of integration concerns space, in the second, time. Only if face-processing relies on dedicated, specialized processing nodes, these two types of integration would be hypothesized to occur in and affect processing nodes related to face-processing. The present thesis asks exactly this: how does spatial and temporal integration affect processing in human face regions?

Part one: Chapter two

As illustrated in the introduction, it is not fully understood how faces are represented in the neuronal core network of face processing. Its detailed involvement in identity and expressions processing and how it represents this facial information remains still controversial. In studies using tasks based on identity recognition, the FFA has been found to contain whole face as well as part based representations (Rossion et al., 2000; Yovel and Kanwisher, 2004;

Rotshtein et al., 2007; Schiltz and Rossion, 2006; Schiltz et al., 2010; Liu et al., 2010; Golarai et al., 2015). The OFA on the other hand shows stronger part-based processing with holistic representations only to a smaller extend (Liu et al., 2010; Golarai et al., 2015; Schiltz and Rossion, 2006). The STS has hardly been included in studies on identity representation due to the fact that it has not been considered to be involved in identity recognition at all (Rhodes et al., 2009). Studies on the representation of expressions in the STS have reported that its activity pattern closely resemble the perceptual input (Said et al., 2010). This suggests a rather part-based representation in the STS which would fit to the rather part-based processing strategies described for facial expressions.

The study presented in Chapter two of this thesis therefore aims to systematically investigate part-based and whole face representations of facial expressions in the three areas of the core face network (FFA, OFA and STS). During a fMRI-experiment, whole faces and facial parts (mouth and eyes) of fearful and happy expressions were presented to the participants. We then measured the similarity of activity patterns in OFA, FFA and STS between whole faces and facial parts across and within the two expressions. In detail, we tried to answer four different questions.

First, we examined if we could classify the facial expressions from either full faces, the eye-region or the mouth. Classification from the mouth was successful in the OFA and STS, with stronger effects in the left hemisphere and a trend was observed in the FFA.

Secondly, we tested if the average pattern of the eyes and the mouth of a given expression could predict the pattern of an emotionally congruent whole face. This test paralleled an approach from a paper of MacEvoy and Epstein (2011) who predicted scene pattern from averaged pattern of their constituting objects. If prediction was successful, we concluded that the ROI encoded facial expressions both in whole faces and facial parts. We found significant effects in the STS and again a trend in the FFA.

In a third step, we wanted to find out, if an area stronger represents whole faces irrespectively from the presented expressions, potentially indicating holistic processes; or if an area rather processes the facial expression, independently if a whole face or facial parts were shown. This approaches balances low-level confounds of different stimulus sizes of whole faces and facial parts against high level confounds of differences between facial expressions. While the FFA rather represents whole faces independently from the facial expressions, the STS shows the opposite pattern.

Our fourth question investigated, if the differential impact of the eyes- and mouth-region that is behaviorally described for the recognition of facial expressions (Calvo and Nummenmaa, 2008; Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011) could also be found in the underlying neuronal patterns of emotional faces and facial parts. The OFA seems to represent eyes and mouth equally independent from the expression type. In the FFA, in contrast, the eyes seems to have a higher informational content. The STS, however, represents eyes and mouth differentially depending on the expression type.

Taken together, our results extend the knowlegde of how expressional faces and facial parts are represented in the brain. It supports the proposed functional distinction of areas of the core network with the OFA being rather an entry point for facial information into the core system (Pitcher et al., 2007; Pitcher et al., 2011), the FFA being mainly involved in whole face representations and idenity recognition and the STS playing a major role in the representation of facial expression. However, they also adds evidence that this distinction is less strict as previously assumed by also showing specific effects of expressional information representation in the FFA.

Part two: Chapter three to five

The second part of this thesis examines temporal information in dynamic faces. For decades, static faces have been used for research, fully neglecting the possibly rich information embodied in facial dynamics. In daily life, we usually encounter moving faces and we easily extract the huge amount of information from those faces within a fraction of seconds. How we do that and if we indeed need all these dynamic information is still unclear. Using moving faces, the recognition of facial expressions improves and motion even enables us to recognize the identity of a face (O'Toole et al., 2002; Thornton and Kourtzi, 2002). However, it is important that this dynamic information is perceived as coherent and reasonable motion. Scrambling the frames of face movies diminishes expression recognition. Also the widely found increase in activity of face core network is absent for scrambled movies. Many studies used morphed dynamic face stimuli instead of real recordings, as they are easier to manipulate and better to control for low-level aspects. However, morphs often contain only linear changes of facial features which do not resemble the natural time course of real facial dynamics. Curio et al., (2006) had shown that participants are indeed sensitive to such small differences in facial dynamics. In our experiment we decided to use genuine recordings of facial expressions. We introduced a 2x2-factorial design to manipulate the timeline of expression unfolding as well as the ecological meaning of the presented expression in a highly controlled fashion. We used recordings of increasing and decreasing facial expressions (factor “emotion-direction”), which we played in forward as well as reversed frame order (factor “timeline”). This approach allowed us control for low-level visual properties and match the static content and motion energy within each factor. We assume differences in the spatio-temporal course of increasing and decreasing expressions, so reversing the timeline is different from changing the emotion direction.

The first (Chapter three) and the second (Chapter four) experiment included a stimulus set of fearful expressions. The first study (Chapter three) describes a fMRI-experiment to investigate the response of the core network of face processing towards our manipulation. We examined the responses of OFA, FFA and STS towards the two factor “timeline” and “emotion-direction”. The occipital face area (OFA) was sensitive to the factor “timeline”. The responses in FFA were different between the two levels of the factor “emotion-direction”, however, these differences were timeline-dependent. Differential responses between increasing and decreasing expressions occurred only within the natural frame order. The STS was sensitive to the factor “timeline”, which, however, was emotion-direction-dependent as it only occurred for decreasing fear. In summary, we find an interactive mechanism sensitive to both, the ecological meaning as well as the prototypical unfolding of facial dynamics.

The second study (Chapter four) provides a perceptual evaluation of the same stimulus set. Participants were instructed to evaluate the stimuli regarding three aspects: 1. the intensity of the presented emotion, 2. the artificialness of the actor movement, which should be evaluated independent from the quality of the expression portrayal and 3. the convincingness or recognizability of the presented expression, that is, how well the expression was portrayed by the actor. We find all three ratings to be influenced by both factors “timeline” and “emotion-direction”. Reversing the frame order and thereby altering the movement trajectories of the facial muscles involved in fearful expressions results lower ratings for perceived emotional intensity and convincingness but higher ratings of perceived artificialness. Manipulating dynamic information thereby not only affects processing on the neuronal level but also changes our perceptual evaluation of the facial expressions.

The third study (Chapter five) then introduced a new stimulus set, comprising additional expressions to test the generalizability of our previous findings. It also included a group of participants with ASD to compare differences in the perception of our stimuli between a group of TD and a group of ASD participants with impairments in face processing. The task was to identify the presented emotion type and to evaluate the intensity and authenticity of each expressions. Additionally, participants had to decide on the timeline of the stimuli. We find expression recognition influenced by our factor timeline with surprisingly lower expressions recognition rates in forward movies. Expression recognitions rates additionally were lower in ASD participants and this effect was stronger for negative emotions. Another main effect of group was found for the autenticity rating with ASD participants rating stimuli as less authentic than the control group. Both groups rated increasing expressions as more a authentic and intense than decreasing expressions, but this was true for intensity only for the natural frame order. The recognition of the presented timeline was influenced by both factors but did not reveal any group difference. From our results we conclude, that participants with and without ASD are influenced by the interacting dynamics of facial trajectories and changes of expression intensity. However, our findings also indicate differences how both groups percieve dynamic faces. Anyhow, these differences seem not to be the specific trajectories of facial dynamics.

1.7 General Discussion

This thesis targets two highly topical questions in the research of face processing concerning the spatial and temporal integration of facial features. On a behavioral level, we intensively investigated the impact of temporal information in TD subjects and further in a group of participants with ASD, as the latter are reported to have impairments in social interaction and face processing. Additionally we investigated the neural responses of the core face processing regions in the binding of facial information over time and space.

Representation of facial expressions

First, we explored how facial expressions are represented in areas dedicated to face processing using a novel approach based on pattern similarity. It is well known how identity is represented, but already behavioral findings point to differences in the way how facial expressions are processed (McKelvie, 1995; Prkachin, 2003; Calvo and Nummenmaa, 2008; Derntl et al., 2009; Narme et al., 2011; Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011; Gosselin and Schyns, 2001; Blais et al., 2012) and differential neural routes have been proposed for the processing of identity and expressions (Haxby et al., 2000). The first Chapter of this thesis provides novel insights on this issue.

Our results extend previous findings on emotion representation in the FFA. The FFA has been shown to repond stronger to emotional compared to neutral faces (Dolan et al., 1996; Gerber et al., 2008; Ishai et al., 2004; Pessoa et al., 2002; Surguladze et al., 2003; Vuilleumier and Pourtois, 2007; Vuilleumier et al., 2001; Winston et al., 2003) but the detailed representation of expressional faces and facial parts had not been investigated yet. We showed that the FFA represents expressional information in facial parts, but still exhibits a strong full face representation that is emotion-independent. The findings that the eyes are by trend stronger represented than the mouth is in line with the importance of the eye-region for identity recognition (McKelvie, 1976; Gilad et al., 2009; Sormaz et al., 2013; Malcolm et al., 2008).

The OFA represents facial parts independently of or invariant with regards to the emotional content, supporting the view that it represents an early stage in the representation of facial information (Pitcher et al., 2007; Pitcher et al., 2011). Pitcher et al., (2007) showed that early, but not late disruption of OFA function by TMS pulses impaired identity discrimination tasks. Kadosh et al., 2011 extended the findings on also mid-latency processing stages for identity as well expressional information. From our design and experimental method, however, we can not say whether the activation in OFA precedes activation in FFA and STS or not.

The STS, usually neglected in studies on identity processing, shows a stronger part-based representation of facial expressions. This could be hypothesized from previous studies (Said et al., 2010) and adds further evidence to an involvement of the STS in expression recognition.

Our results add evidence to the ‘bias rather than […] categorical dissociation’ of identity and expression recognition in FFA and STS, respectively, as proposed by Calder & Young, (2005, p.464). We usually have to extract various types of information from faces (i.e. gender, identity, expression, trustworthiness, attractiveness) in parallel and we do that effortlessly. It is therefore not surprinsing that this information at least somewhat overlaps in the neural face processing system. Our results nevertheless support some functional distinction between the areas of the core network of face processing that has been described in the existing literature.

Temporal information of dynamic facial expressions

The second chapter addresses another timely question in todays' face research. Dynamic face stimuli have been described as more ecologically valid, and they have been found to enhance different aspects in face perception and processing compared to static faces (Ambadar et al., 2005; Wehrle et al., 2000; O'Toole et al., 2002; Thornton and Kourtzi, 2002; Hill and Johnston, 2001). It is, however, still unclear what actually drives this dynamic advantage. One difficulty in studying the impact of facial dynamics is to come up with proper control conditions. This thesis introduces a new design to manipulate temporal information in facial expressions. Using a well-controlled 2x2 design we manipulated the emotional and temporal information of genuine facial movement while keeping all low level aspects in the stimuli constant. We conducted three studies to extensively investigate the effects of our manipulation.

We were able to show that the core network of face processing is sensitive to timeline manipulations (Chapter two), even if subjects were informed of the actual manipulation (i.e. reversal of time). We describe a sensitivity for the socially relevant information of the emotional state of another person (factor "emotion-direction") in the FFA, which, however, was was timeline-dependent. This findings underline importance of using genuine recordings of facial expressions in research. Often, morph stimuli have been used to study dynamic face perception as they allow for more precise control of low-level stimulus features. However, if morphs are not created from real movement trajectories of facial muscles, the evaluation of the stimuli was found to be different (Curio et al., 2006; Dobs et al., 2014) and we extend this findings with effects found on the neuronal processing level. The STS responded differentially towards the timeline manipulation, however here we find an emotion-direction-dependence. Models on the processing of biological motion propose motion pattern neurons that integrate information from so-called snaphot-neurons as well as optic-flow-detector neurons (Giese and Poggio, 2003). Such circuits would be expected to be sensitive to temporal sequences of biological motion. Electrophysiology has detected “snapshot” neurons that respond to static information as well as “motion” neurons responsive to motion trajectories in the STS while viewing body actions (Vangeneugden et al., 2009). Our study provides first fMRI-evidence of potential predictive coding in the STS for processing dynamic facial information. The OFA, in contrast, only showed sensitivity to the factor timeline. There is evidence that dynamic faces enhance the coupling of OFA and STS (Foley et al., 2012). The OFA additionally has been shown to act as an entry point of facial information into the core system of face processing (Pitcher et al., 2007; Pitcher et al., 2011). Our findings support this role of a early representation of facial information and its connection to the STS, even if we did not explicitly tested for congruent actitivty pattern.

Furthermore, the perceptual evaluation of our stimuli is influenced by the induced subtle changes in temporal information of our stimulus manipulation (Chapter three). Ratings of the emotional intensity, the artificialness of the facial movement and the convincingness or plausibility of the emotion portrayal were affected by the timeline manipulation. In general, reversed movies were described as less intense, less convincing and more artificial. Our results provides evidence that our visual system is sensitive to the temporal order of facial muscle movement, even if we are not aware that their motion trajectories have been artificially manipulated. This might be mediated by an automatic process called facial mimicry, mediated by the neuron mirron system. When watching facial expressions of other people, subtle activations in facial muscles can be detected (Hoffman, 1984; Hatfield et al., 1993) and are thought to enhance our understanding of the presented information. If genuine recordings of facial expressions are reversed, the order of muscle activation does not follow the usual trajectories anymore and this might interfere with the automatic facial reactions elicited in the viewer.

As our first two studies included only fearful expressions, we wanted to extend our findings to other basic expressions and test the generalizability of our results. We introduced a new stimulus (Chapter five) containing four additional facial expressions: surprise, anger, happiness and sadness. We further included a group of participants with ASD, as they have been shown to have difficulties in the perception and evaluation of faces and facial expressions. Using slightly different behavioral ratings we aimed to extend our previous results. Again, we find perceptual parameters like intensity and authenticity to be affected by emotion- and timeline-direction. Additionally, we find group differences between TD and ASD participants for expression recognition rates and authenticity ratings. The results show that in both TD and ASD groups, facial trajectory dynamics and the direction of expression change interact on their influence of face perception. They also confirm generic differences between TD and ASD in percieving dynamic faces, but their don´t seem to be specific to facial trajectory dynamics. As studies of the ASD however indicated disturbances in the mirror neuron system and for facial mimicry, we would have expected differential effects for our timeline manipulation between the ASD and the TD group. However, we do not find significant interactions between the factors “timeline” and “group”. This seems surprising at first. As our approach is based on movement asymmetries between increase and decrease of a facial expression, it is likely that those asymmetries differ between the different basic emotions. Additionally, our results show that impairments in the recognition accuracy of facial expression are at least partly affected by the type of facial expression in the ASD group. Averaging across all expressions could therefore mask effects that might be found for single expressions. However, as a limitation of our study, we did not include enough stimuli per expression to consider every expression individually. In a next step, we therefore plan to extend the motion analysis from Chapter four on the stimuli used in Chapter five. This might enable us to split the data into high- and low-asymmetric expressions, thereby increasing the number of stimuli per group and repeat the analysis. This might be a more sensitive way to approach the data. It would be of great interest to find out whether ASD is affected by subtle temporal changes in facial muscle movement as this would point towards a better face perception as often is assumed. Of course, for the interpretation of such results, one has to consider that our ASD group mainly consists of adult participants diagnosed with Asperger Syndrome. Although impairments in face recognition persist a lifetime, they still might have developed alternative strategies to navigate in real life and social interactions. Asperger and autistic patients have also been shown to be differentially affected regarding motion perception. Potential findings in further analysis therefore should be considered with care as they might have a limited generalizability and rather describe effects for a certain portion of people from the ASD.

1.8 Conclusion

This thesis targets a very important issue in today´s face research: how are facial features integrated over space and over time. In four experiments this questions has been approached on a behavioral and neuronal level. Faces have been shown to be represented differentially than other objects, as their features are often combined to a so-called ‘holistic representation’. How exactly this integration of facial features over space is done, and if there are differences depending on which information, i.e. identity or facial expression, is represented from the face remains controversal. Previous findings indicate differences in the holistic or part-based representation of facial identity and facial expressions, respectively, on the behaviorally and on the neuronal level. This questions is targeted by the first part of the thesis. Our findings supports the differential representation of facial information in the areas of the core network of face processing, supporting the proposed division of identity and expressions. Albeit it points towards a less strict distinction in the brain than proposed by Haxby et al., (2000), underlining the interaction of the core areas in extracting various distinct types of information from faces that we see around us. As the debate on the influence of featural, configural and holistic information on the representation of facial information has mainly been approached using static stimuli, it will be very interesting to expand previous research with the use of dynamic face stimuli, providing more ecologically valid stimuli, in the future. The second part of the thesis however points towards a very important aspect when choosing such stimuli. It examines another important question that was raised within the last couple of years: How is the temporal information in dynamic faces integrated over time and how does it affect our perception? We used genuine recordings of facial expressions that we manipulated in their fine temporal trajectories. We extend previously shown sensitivities of subjects to subtle dynamic information (Curio et al., 2006; Dobs et al., 2014) not only behaviorally but now also on an neuronal level that has not been investigated before. Our neuronal face processing system is tuned to fine temporal information. Even though we are able to extract basic information despite subtle changes in facial motion trajectories, these subtle changes are still reflected in the activation of the core network of face processing and they affect our higher-level perceptual evaluation of the faces we see. For future face research it therefore seems to be important to use genuine recordings of faces or, at least, realistically animated facial morphs not to miss important details to study the representation of facial information in the brain and their evaluation on a behavioral level.

1.9 References

Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P., Damasio, A.R., 2005. A mechanism for impaired fear recognition after amygdala damage. Nature 433, 68–72.

Ambadar, Z., Schooler, J.W., Cohn, J.F., 2005. Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. Psychological science : a journal of the American Psychological Society / APS 16, 403–410.

Amishav, R., Kimchi, R., 2010. Perceptual integrality of componential and configural information in faces. Psychon Bull Rev 17, 743–748.

Annaz, D., Campbell, R., Coleman, M., Milne, E., Swettenham, J., 2012. Young children with autism spectrum disorder do not preferentially attend to biological motion. J Autism Dev Disord 42, 401–408.

Ashwin, C., Baron-Cohen, S., Wheelwright, S., O'Riordan, M., Bullmore, E.T., 2007. Differential activation of the amygdala and the "social brain" during fearful face-processing in Asperger Syndrome. Neuropsychologia 45, 2–14.

Atkinson, A.P., 2009. Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. Neuropsychologia 47, 3023–3029.

Bal, E., Harden, E., Lamb, D., Van Hecke, A.V., Denver, J.W., Porges, S.W., 2010. Emotion recognition in children with autism spectrum disorders: relations to eye gaze and autonomic state. J Autism Dev Disord 40, 358–370.

Barnea-Goraly, N., Kwon, H., Menon, V., Eliez, S., Lotspeich, L., Reiss, A.L., 2004. White matter structure in autism: preliminary evidence from diffusion tensor imaging. Biol Psychiatry 55, 323–326.

Bartlett, J. C., Searcy, J. H., & Abdi, H., 2003. What are the routes to face recognition? In M. Peterson & G. Rhodes (Eds.), Perception of faces, objects, and scenes: Analytic and holistic processes (pp. 21–52). Oxford: Oxford University Press.

Belmonte, M.K., Allen, G., Beckel-Mitchener, A., Boulanger, L.M., Carper, R.A., Webb, S.J., 2004. Autism and abnormal development of brain connectivity. Journal of Neuroscience 24, 9228–9231.

Bird, G., Catmur, C., Silani, G., Frith, C., Frith, U., 2006. Attention does not modulate neural responses to social stimuli in autism spectrum disorders. NeuroImage 31, 1614–1624.

Blais, C., Roy, C., Fiset, D., Arguin, M., Gosselin, F., 2012. The eyes are not the window to basic emotions. Neuropsychologia 50, 2830–2838.

Blake, R., Turner, LM., Smoski, MJ., Pozdol, SL., Stone, WL., 2003. Visual recognition of biological motion is impaired in children with autism. Psychol Sci 14: 151–157.

Blake, R., Shiffrar, M., 2007. Perception of human motion. Annu. Rev. Psychol. 58, 47–73.

Bodamer J. (1947) Die Prosop-Agnosie, Arch. Pslxhirrr. Nercrnkr.. 179: 6654.

Bonda, E., Petrides, M., Ostry, D., Evans, A., 1996. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. J Neurosci 16, 3737–3744.

Bruce, V., Young, A., 1986. Understanding face recognition. British Journal of Psychology 77, 305–327.

Calder, A.J., Young, A.W., 2005. Understanding the recognition of facial identity and facial expression. Nat Rev Neurosci 6, 641–651.

Calder, A.J., Young, A.W., Keane, J., Dean, M., 2000. Configural information in facial expression perception. J Exp Psychol Hum Percept Perform 26, 527–551.

Calvo, M.G., Beltrán, D., 2014. Brain lateralization of holistic versus analytic processing of emotional facial expressions. NeuroImage 92, 237–247.

Calvo, M.G., Nummenmaa, L., 2008. Detection of emotional faces: salient physical features guide effective visual search. J Exp Psychol Gen 137, 471–494.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M.J., David, A.S., 2001. Cortical substrates for the perception of face actions:

an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). Brain research Cognitive brain research 12, 233–243.

Carey, S., 1981. The development of face perception. In G. Davies, H.Elli, & J. Shepard (Eds), Perceiving and remembering faces (pp. 12-24). London: Academic Press.

Castelli, F., 2005. Understanding emotions from standardized facial expressions in autism and normal development. Autism 9, 428–449.

Corbett, B.A., Carmean, V., Ravizza, S., Wendelken, C., Henry, M.L., Carter, C., Rivera, S.M., 2009. A functional and structural study of emotion and face processing in children with autism. Psychiatry Res 173, 196–205.

Critchley, H.D., Daly, E.M., Bullmore, E.T., Williams, S.C., Van Amelsvoort, T., Robertson, D.M., Rowe, A., Phillips, M., McAlonan, G., Howlin, P., Murphy, D.G., 2000. The functional neuroanatomy of social behaviour: changes in cerebral blood flow when people with autistic disorder process facial expressions. Brain 123 (Pt 11), 2203–2212.

Curio, C., Breidt, M., Kleiner, M., Vuong, Q.C., Giese, M.A., Bülthoff, H.H., 2006. Semantic 3D motion retargeting for facial animation, the 3rd symposium. ACM, New York, New York, USA.

Dalton, K.M., Nacewicz, B.M., Johnstone, T., Schaefer, H.S., Gernsbacher, M.A., Goldsmith, H.H., Alexander, A.L., Davidson, R.J., 2005. Gaze fixation and the neural circuitry of face processing in autism. Nat Neurosci 8, 519–526.

Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y., Iacoboni, M., 2006. Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. Nat Neurosci 9, 28–30.

Decety, J., 1999. Neural mechanisms subserving the perception of human actions. Trends Cogn Sci (Regul Ed) 3, 172–178.

Deeley, Q., Daly, E.M., Surguladze, S., Page, L., Toal, F., Robertson, D., Curran, S., Giampietro, V., Seal, M., Brammer, M.J., Andrew, C., Murphy, K., Phillips, M.L., Murphy, D.G.M., 2007. An event related functional magnetic resonance imaging study of facial emotion processing in Asperger syndrome. Biol Psychiatry 62, 207–217.

Derntl, B., Habel, U., Windischberger, C., Robinson, S., Kryspin-Exner, I., Gur, R.C., Moser, E., 2009. General and specific responsiveness of the amygdala during explicit emotion recognition in females and males. BMC Neurosci 10, 91–14.

Dobs, K., Bülthoff, I., Breidt, M., Vuong, Q.C., Curio, C., Schultz, J., 2014. Quantifying human sensitivity to spatio-temporal information in dynamic faces. Vision Res 100, 78–87.

Dolan, R.J., Fletcher, P., Morris, J., Kapur, N., Deakin, J.F., Frith, C.D., 1996. Neural activation during covert processing of positive emotional facial expressions. NeuroImage 4, 194–200.

Domes, G., Heinrichs, M., Kumbier, E., Grossmann, A., Hauenstein, K., Herpertz, S.C., 2013. Effects of intranasal oxytocin on the neural basis of face processing in autism spectrum disorder. Biol Psychiatry 74, 164–171.

Eisenbarth, H., Alpers, G.W., 2011. Happy mouth and sad eyes: scanning emotional facial expressions. Emotion 11, 860–865.

Ellis, H.D., Florence, M., 1990a. Bodamer's (1947) paper on prosopagnosia. Cognitive Neuropsychology 7, 81–105.

Ellis, H.D., 1990. Developmental trends in face recognition. The Psychologist, 3, 124-129.

Ekman, P., Friesen, W.V., Hager, J.C., 2002. The Facial Action Coding System, 2nd ed. Research Nexus eBook, Salt Lace City, UT.

Farah, M.J., Wilson, K.D., Drain, M., Tanaka, J.N., 1998. What is "special" about face perception? Psychol Rev 105, 482–498.

Flack, T.R., Andrews, T.J., Hymers, M., Al-Mosaiwi, M., Marsden, S.P., Strachan, J.W.A., Trakulpipat, C., Wang, L., Wu, T., Young, A.W., 2015. Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. CORTEX 69, 14–23.

Foley, E., Rippon, G., Thai, N.J., Longe, O., Senior, C., 2012. Dynamic facial expressions evoke distinct activation in the face perception network: a connectivity analysis study. J Cogn Neurosci 24, 507–520.

Fox, C.J., Iaria, G., Barton, J.J.S., 2009. Defining the face processing network: Optimization of the functional localizer in fMRI. Hum Brain Mapp 30, 1637–1651.

Freitag, C.M., Konrad, C., Häberlen, M., Kleser, C., Gontard, von, A., Reith, W., Troje, N.F., Krick, C., 2008. Perception of biological motion in autism spectrum disorders. Neuropsychologia 46, 1480–1494.

Furl, N., Garrido, L., Dolan, R.J., Driver, J., Duchaine, B., 2011. Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. J Cogn Neurosci 23, 1723–1740.

Furl, N., van Rijsbergen, N.J., Kiebel, S.J., Friston, K.J., Treves, A., Dolan, R.J., 2010. Modulation of Perception and Brain Activity by Predictable Trajectories of Facial Expressions. Cerebral Cortex 20, 694–703.

Galton, F (1879) Composite portraits, made by combining those of many different persons into a single, resultant figure. Journal of the Anthropological Institute, 8, 132-144

Gamer, M., Büchel, C., 2009. Amygdala activation predicts gaze toward fearful eyes. Journal of Neuroscience 29, 9123–9126.

Gamer, M., Schmitz, A.K., Tittgemeyer, M., Schilbach, L., 2013. The human amygdala drives reflexive orienting towards facial features. Current Biology 23, R917–R918.

George, M.S., Ketter, T.A., Gill, D.S., Haxby, J.V., Ungerleider, L.G., Herscovitch, P., Post, R.M., 1993. Brain regions involved in recognizing facial emotion or identity: an oxygen-15 PET study. J Neuropsychiatry Clin Neurosci 5, 384–394.

Gepner, B., Deruelle, C., Grynfeltt, S., 2001. Motion and emotion: a novel approach to the study of face processing by young autistic children. J Autism Dev Disord 31, 37–45.

Gerber, A.J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J., Peterson, B.S., 2008. An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. Neuropsychologia 46, 2129–2139.

Giese, M.A., Poggio, T., 2003. Cognitive neuroscience: Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci 4, 179–192.

Gilad, S., Meng, M., Sinha, P., 2009. Role of ordinal contrast relationships in face encoding. Proceedings of the National Academy of Sciences 106, 5353–5358.

Golarai, G., Ghahremani, D.G., Eberhardt, J.L., Gabrieli, J.D.E., 2015. Distinct representations of configural and part information across multiple face-selective regions of the human brain. Front. Psychology 6, 1710.

Gosselin, F., Schyns, P.G., 2001. Bubbles: a technique to reveal the use of information in recognition tasks. Vision Res 41, 2261–2271.

Gosselin, F., Spezio, M.L., Tranel, D., Adolphs, R., 2011. Asymmetrical use of eye information from faces following unilateral amygdala damage. Social Cognitive and Affective Neuroscience 6, 330–337.

Grossman, R.B., Tager-Flusberg, H., 2008. Reading faces for information about words and emotions in adolescents with autism. Research in Autism Spectrum Disorders 2, 681–695.

Hadjikhani, N., Joseph, R.M., Snyder, J., Chabris, C.F., Clark, J., Steele, S., McGrath, L., Vangel, M., Aharon, I., Feczko, E., Harris, G.J., Tager-Flusberg, H., 2004. Activation of the fusiform gyrus when individuals with autism spectrum disorder view faces. NeuroImage 22, 1141–1150.

Hadjikhani, N., Joseph, R.M., Snyder, J., Tager-Flusberg, H., 2007. Abnormal activation of the social brain during face perception in autism. Hum Brain Mapp 28, 441–449.

Hatfield, E., Cacioppo, J. T., and Rapson, R. L., 1993. Emotional contagion. Curr. Dir. Psychol. Sci. 2, 96–99.

Haxby, J., Hoffman, E., Gobbini, M., 2000. The distributed human neural system for face perception. Trends Cogn Sci (Regul Ed) 4, 223–233.

Hill, H., Johnston, A., 2001. Categorizing sex and identity from the biological motion of faces. Curr Biol 11, 880–885.

Hobson, R.P., 1986a. The autistic child's appraisal of expressions of emotion. J Child Psychol Psychiatry 27, 321–342.

Hobson, R.P., 1986b. The autistic child's appraisal of expressions of emotion: a further study. J Child Psychol Psychiatry 27, 671–680.

Hobson, R.P., Ouston, J., Lee, A., 1988. Emotion recognition in autism: coordinating faces and voices. Psychol Med 18, 911–923.

Hoffman, M. L., 1984. “Interaction of affect and cognition in empathy,” in Emotions, Cognition, and Behavior, eds C. E. Izard, J. Kagan, and R. B. Zajonc (Cambridge: Cambridge University Press), 103–131.

Hoffman, E.A., Haxby, J.V., 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nat Neurosci 3, 80–84.

Horwitz, B., Rumsey, J.M., Grady, C.L., Rapoport, S.I., 1988. The cerebral metabolic landscape in autism. Intercorrelations of regional glucose utilization. Arch. Neurol. 45, 749–755.

Hubert, B., Wicker, B., Moore, D.G., Monfardini, E., Duverger, H., Da Fonséca, D., Deruelle, C., 2007. Brief report: recognition of emotional and non-emotional biological motion in individuals with autistic spectrum disorders. J Autism Dev Disord 37, 1386–1392.

Hubl, D., Bölte, S., Feineis-Matthews, S., Lanfermann, H., Federspiel, A., Strik, W., Poustka, F., Dierks, T., 2003. Functional imbalance of visual pathways indicates alternative face processing strategies in autism. Neurology 61, 1232–1237.

Humphreys, K., Hasson, U., Avidan, G., Minshew, N., Behrmann, M., 2008. Cortical patterns of category-selective activation for faces, places and objects in adults with autism. Autism Res 1, 52–63.

Ishai, A., Pessoa, L., Bikle, P.C., Ungerleider, L.G., 2004. Repetition suppression of faces is modulated by emotion. Proc Natl Acad Sci USA 101, 9827–9832.

Jones, C.R.G., Swettenham, J., Charman, T., Marsden, A.J.S., Tregay, J., Baird, G., Simonoff, E., Happé, F., 2011. No evidence for a fundamental visual motion processing deficit in adolescents with autism spectrum disorders. Autism Res 4, 347–357.

Kadosh, K.C., Walsh, V., Kadosh, R.C., 2011. Investigating face-property specific processing in the right OFA. Social Cognitive and Affective Neuroscience 6, 58–65.

Kaiser, M.D., Pelphrey, K.A., 2012. Disrupted action perception in autism: Behavioral evidence, neuroendophenotypes, and diagnostic utility. Accident Analysis and Prevention 2, 25–35.

Kilts, C.D., Egan, G., Gideon, D.A., Ely, T.D., Hoffman, J.M., 2003. Dissociable Neural Pathways Are Involved in the Recognition of Emotion in Static and Dynamic Facial Expressions. NeuroImage 18, 156–168.

Kimchi, R., Amishav, R., 2010. Faces as perceptual wholes: The interplay between component and configural properties in face processing. Visual Cognition 18, 1034–1062.

Kleinhans, N.M., Johnson, L.C., Richards, T., 2009. Reduced neural habituation in the amygdala and social impairments in autism spectrum disorders. American Journal of Psychiatry 166, 467-475.

Kleinhans, N.M., Richards, T., Weaver, K., Johnson, L.C., Greenson, J., Dawson, G., Aylward, E., 2010. Association between amygdala response to emotional faces and social anxiety in autism spectrum disorders. Neuropsychologia 48, 3665–3670.

Koshino, H., Kana, R.K., Keller, T.A., Cherkassky, V.L., Minshew, N.J., Just, M.A., 2008. fMRI investigation of working memory for faces in autism: visual coding and underconnectivity with frontal areas. Cerebral Cortex 18, 289–300.

Labar, K.S., Crupain, M.J., Voyvodic, J.T., Mccarthy, G., 2003. Dynamic perception of facial affect and identity in the human brain. Cereb Cortex 13, 1023–1033.

Leder, H., Bruce, V., 2000. When inverted faces are recognized: the role of configural information in face recognition. Q J Exp Psychol A 53, 513–536.

Lindner, J.L., Rosén, L.A., 2006. Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. J Autism Dev Disord 36, 769–777.

Liu, J., Harris, A., Kanwisher, N., 2010. Perception of face parts and face configurations: an FMRI study. J Cogn Neurosci 22, 203–211.

Loveland, K.A., Tunali-Kotoski, B., Chen, Y.R., Ortegon, J., Pearson, D.A., Brelsford, K.A., Gibbs, M.C., 1997. Emotion recognition in autism: verbal and nonverbal information. Dev Psychopathol 9, 579–593.

Malcolm, G.L., Lanyon, L.J., Fugard, A.J.B., Barton, J.J.S., 2008. Scan patterns during the processing of facial expression versus identity: an exploration of task-driven and stimulus-driven effects. JOV 8, 2.1–9.

Malisza, K.L., Clancy, C., Shiloff, D., Holden, J., Jones, C., Paulson, K., Yu, D.C.T., Summers, R., Chudley, A.E., 2011. Functional magnetic resonance imaging of facial information processing in children with autistic disorder, attention deficit hyperactivity disorder and typically developing controls. Int J Adolesc Med Health 23, 269–277.

Maurer, D., Grand, R.L., Mondloch, C.J., 2002. The many faces of configural processing. Trends Cogn Sci (Regul Ed) 6, 255–260.

Maximo, J.O., Cadena, E.J., Kana, R.K., 2014. The implications of brain connectivity in the neuropsychology of autism. Neuropsychology review 24, 16–31. doi:10.1007/s11065-014-9250-0

McKelvie, S.J., 1995. Emotional expression in upside-down faces: evidence for configurational and componential processing. Br J Soc Psychol 34 (Pt 3), 325–334.

McKelvie, S.J., 1976. The role of eyes and mouth in the memory of a face. The American Journal of Psychology 89, 311–323.

Monk, C.S., Weng, S.-J., Wiggins, J.L., Kurapati, N., Louro, H.M.C., Carrasco, M., Maslowsky, J., Risi, S., Lord, C., 2010. Neural circuitry of emotional face processing in autism spectrum disorders. J Psychiatry Neurosci 35, 105–114.

Moore DG., Hobson RP, Lee A., 1997. Components of person perception: An investigation with autistic, non-autistic retarded and typically developing children and adolescents. Br J Develop Psych 15: 401–423.

Murphy, E.R., Foss-Feig, J., Kenworthy, L., Gaillard, W.D., Vaidya, C.J., 2012. Atypical Functional Connectivity of the Amygdala in Childhood Autism Spectrum Disorders during Spontaneous Attention to Eye-Gaze. Autism Res Treat 2012, 652408–12.

Murphy, P., Brady, N., Fitzgerald, M., Troje, N.F., 2009. No evidence for impaired perception of biological motion in adults with autistic spectrum disorders. Neuropsychologia 47, 3225–3235.

Narme, P., Bonnet, A.-M., Dubois, B., Chaby, L., 2011. Understanding facial emotion perception in Parkinson's disease: the role of configural processing. Neuropsychologia 49, 3295–3302.

Nestor, A., Plaut, D.C., Behrmann, M., 2011. Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. Proceedings of the National Academy of Sciences 108, 9998–10003.

O'Toole, A.J., Roark, D.A., Abdi, H., 2002. Recognizing moving faces: a psychological and neural synthesis. Trends Cogn Sci (Regul Ed) 6, 261–266.

Parron, C., Da Fonseca, D., Santos, A., Moore, D.G., Monfardini, E., Deruelle, C., 2008. Recognition of biological motion in children with autistic spectrum disorders. Autism 12, 261–274.

Peelen, M.V., Wiggett, A.J., Downing, P.E., 2006. Patterns of fMRI Activity Dissociate Overlapping Functional Brain Areas that Respond to Biological Motion. Neuron 49, 815–822.

Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G., 2002. Neural processing of emotional faces requires attention. Proc Natl Acad Sci USA 99, 11458–11463.

Pierce, K., Müller, R.A., Ambrose, J., Allen, G., Courchesne, E., 2001. Face processing occurs outside the fusiform “face area” in autism: evidence from functional MRI. Brain 124, 2059–2073.

Pierce, K., Redcay, E., 2008. Fusiform function in children with an autism spectrum disorder is a matter of "who". Biol Psychiatry 64, 552–560.

Piggot, J., Kwon, H., Mobbs, D., Blasey, C., Lotspeich, L., Menon, V., Bookheimer, S., Reiss, A.L., 2004. Emotional attribution in high-functioning individuals with autistic spectrum disorder: a functional imaging study. J Am Acad Child Adolesc Psychiatry 43, 473–480.

Pitcher, D., Dilks, D.D., Saxe, R.R., Triantafyllou, C., Kanwisher, N., 2011a. Differential selectivity for dynamic versus static information in face-selective cortical regions. NeuroImage 56, 2356–2363.

Pitcher, D., Walsh, V., Duchaine, B., 2011b. The role of the occipital face area in the cortical face perception network. Exp Brain Res 209, 481–493.

Prkachin, G.C., 2003. The effects of orientation on detection and identification of facial expressions of emotion. Br J Psychol 94, 45–62.

Puce, A., Allison, T., Bentin, S., Gore, J.C., Mccarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. J Neurosci 18, 2188–2199.

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. Eur J Neurosci 30, 721–733.

Richler, J.J., Mack, M.L., Palmeri, T.J., Gauthier, I., 2011a. Inverted faces are (eventually) processed holistically. Vision Res 51, 333–342.

Richler, J.J., Palmeri, T.J., Gauthier, I., 2012. Meanings, mechanisms, and measures of holistic processing. Front. Psychology 3, 553.

Richler, J.J., Wong, Y.K., Gauthier, I., 2011b. Perceptual expertise as a shift from strategic interference to automatic holistic processing. Current directions in ….

Rossion, B., 2008. Picture-plane inversion leads to qualitative changes of face perception. Acta Psychol (Amst) 128, 274–289.

Rossion, B., Dricot, L., Devolder, A., Bodart, J.M., Crommelinck, M., de Gelder, B., Zoontjes, R., 2000. Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. J Cogn Neurosci 12, 793–802.

Rotshtein, P., Geng, J.J., Driver, J., Dolan, R.J., 2007. Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: behavioral and functional magnetic resonance imaging data. J Cogn Neurosci 19, 1435–

Rutherford, M.D., Troje, N.F., 2012. IQ predicts biological motion perception in autism spectrum disorders. J Autism Dev Disord 42, 557–565.

Said, C.P., Moore, C.D., Engell, A.D., Todorov, A., Haxby, J.V., 2010a. Distributed representations of dynamic facial expressions in the superior temporal sulcus. JOV 10, 11.

Sato, W., Kochiyama, T., Yoshikawa, S., 2010. Amygdala activity in response to forward versus backward dynamic facial expressions. Brain Res 1315, 92–99.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., Matsumura, M., 2004. Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. Cognitive Brain Research 20, 81–91.

Schiltz, C., Dricot, L., Goebel, R., Rossion, B., 2010. Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. JOV 10, 25.1–16.

Schiltz, C., Rossion, B., 2006. Faces are represented holistically in the human occipito-temporal cortex. NeuroImage 32, 1385–1394.

Schultz, J., Brockhaus, M., Bulthoff, H.H., Pilz, K.S., 2013. What the Human Brain Likes About Facial Motion. Cereb Cortex 23, 1167–1178.

Schultz, J., Pilz, K.S., 2009. Natural facial motion enhances cortical responses to faces. Exp Brain Res 194, 465–475.

Schultz, R.T., 2005. Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. International Journal of Developmental Neuroscience 23, 125–141.

Schweinberger, S.R., Burton, A.M., Kelly, S.W., 1999. Asymmetric dependencies in perceiving identity and emotion: Experiments with morphed faces. Perception & Psychophysics 61, 1102–1115.

Schweinberger, S.R., Soukup, G.R., 1998. Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. Journal of Experimental Psychology 24, 1748-1765.

Sergent, J., Ohta, S., Macdonald, B., Zuck, E., 2007. Segregated processing of facial identity and emotion in the human brain: A pet study. Visual Cognition.

Smith, M.L., Cottrell, G.W., Gosselin, F., Schyns, P.G., 2005. Transmitting and decoding facial expressions. Psychological science : a journal of the American Psychological Society / APS 16, 184–189.

Soria Bauser, D., Suchan, B., 2015. Is the whole the sum of its parts? Configural processing of headless bodies in the right fusiform gyrus. Behav Brain Res 281, 102–110.

Sormaz, M., Andrews, T.J., Young, A.W., 2013. Contrast negation and the importance of the eye region for holistic representations of facial identity. J Exp Psychol Hum Percept Perform 39, 1667–1677.

Spezio, M.L., Huang, P., Castelli, F., 2007. Amygdala damage impairs eye contact during conversations with real people. The Journal of Neurocscience 27, 3994-3997.

Surguladze, S.A., Brammer, M.J., Young, A.W., Andrew, C., Travis, M.J., Williams, S.C.R., Phillips, M.L., 2003. A preferential increase in the extrastriate response to signals of danger. NeuroImage 19, 1317–1328.

Tanaka, J.W., Farah, M.J., 1993. Parts and wholes in face recognition. Q J Exp Psychol A 46, 225–245.

Tardif, C., Lainé, F., Rodriguez, M., Gepner, B., 2007. Slowing down presentation of facial movements and vocal sounds enhances facial expression recognition and induces facial-vocal imitation in children with autism. J Autism Dev Disord 37, 1469–1484.

Thornton, I.M., Kourtzi, Z., 2002a. A matching advantage for dynamic human faces.

Perception 31, 113–132.

Trautmann, S.A., Fehr, T., Herrmann, M., 2009. Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. Brain Res 1284, 100–115.

Vangeneugden, J., Pollick, F., Vogels, R., 2009. Functional differentiation of macaque visual temporal cortical neurons using a parametric action space. Cerebral Cortex 19, 593–611.

Vuilleumier, P., Armony, J.L., Driver, J., Dolan, R.J., 2001. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30, 829–841.

Vuilleumier, P., Pourtois, G., 2007. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. Neuropsychologia 45, 174–194.

Wang, A.T., Dapretto, M., Hariri, A.R., Sigman, M., Bookheimer, S.Y., 2004. Neural correlates of facial affect processing in children and adolescents with autism spectrum disorder. J Am Acad Child Adolesc Psychiatry 43, 481–490.

Wehrle, T., Kaiser, S., Schmidt, S., Scherer, K.R., 2000. Studying the dynamics of emotional expression using synthesized facial muscle movements. J Pers Soc Psychol 78, 105–119.

Weng, S.-J., Carrasco, M., Swartz, J.R., Wiggins, J.L., Kurapati, N., Liberzon, I., Risi, S., Lord, C., Monk, C.S., 2011. Neural activation to emotional faces in adolescents with autism spectrum disorders. Journal of Child Psychology and Psychiatry 52, 296–305.

Wickelgren, I., 2005. Neurology. Autistic brains out of synch? Science 308, 1856–1858.

Wicker, B., Fonlupt, P., Hubert, B., Tardif, C., Gepner, B., Deruelle, C., 2008. Abnormal cerebral effective connectivity during explicit emotional processing in adults with autism spectrum disorder. Social Cognitive and Affective Neuroscience 3, 135–143.

Winston, J.S., O'Doherty, J., Dolan, R.J., 2003. Common and distinct neural responses during direct and incidental processing of multiple facial emotions. NeuroImage 20, 84–97.

Wong, Y.K., Gauthier, I., 2010. Holistic processing of musical notation: Dissociating failures of selective attention in experts and novices. Cognitive, affective & behavioral neuroscience 10, 541–551.

Wright, K., Kelley, E., Poulin-Dubois, D., 2014. Schematic and realistic biological motion identification in children with high-functioning autism spectrum disorder. Research in Autism Spectrum Disorders 8, 1394–1404.

Yin, R., 1969. Looking at upside-down faces. Journal of Experimental Psychology, 81, 141-145.

Young, A.W., Hellawell, D., Hay, D.C., 1987. Configurational information in face perception. Perception 16, 747–759.

Yovel, G., Kanwisher, N., 2004. Face perception: domain specific, not process specific.

Neuron 44, 889–898.

1.10 Declaration of Contribution

This thesis comprises four manuscripts that are either accepted or in preparation for publication. Details about these manuscripts are presented in the following.

1. Reinl, M., Bartels, A. **Neuronal representation of happy and fearful expressions of whole faces and facial parts** (manuscript in preparation): The idea was developed from a paper of MacEvoy and Epstein (2011) by A.B. and M.R. / M.R. created the stimuli, programmed the experiment and collected the data. Data analyzes was designed by A.B. and M.R. / M.R. wrote the manuskript, which was edited by A.B.

2. Reinl, M., Bartels, A. (2014) **Face processing regions are sensitive to distinct aspects of temporal sequence in facial dynamics.** *Neuroimage 102*, 407–415. A.B. devolped the experimental design. / M.R. created the stimuli, programmed the experiment and collected the data. / Both authors decided on the data analyzes. / M.R. and A.B wrote the manuskript.

3. Reinl, M., & Bartels, A. (2015) **Perception of temporal asymmetries in dynamic facial expressions.** *Frontiers in Psychology, 6*, 1107. A.B. and M.R. devolped the experimental design. / M.R. created the stimuli, programmed the experiment and collected the data. / Both authors analyzed the data. / M.R. wrote the manuskript, which was edited by A.B.

4. Reinl, M., Bartels, A., & Wildgruber D., **The impact of temporal asymmetries on the evaluation of dynamic facial expressions in Autism Spectrum Disorder** (manuscript in preparation): All authors devolped the experimental design. / M.R. created the stimuli, programmed the experiment and collected the data. / All authors decided on the data analyzes / M.R. wrote the manuskript, which was edited by A.B.

Part of this work was also presented at the following conferences:
Reinl M & Bartels A (2011): Dynamic faces: fMRI reveals timeline specific responses to facial expression changes, 41st Annual Meeting of the Society for Neuroscience (Neuroscience 2011), Washington, DC, USA.

2 Neuronal representation of happy and fearful expressions in whole faces and facial parts

2.1 Abstract

This study investigates whole face and part-based representations of facial expressions in face-related brain areas. Studies based on identity recognition describe whole face as well as part-based representations in the fusiform face area (FFA) while the occipital face area (OFA) mainly shows part-based processing. The superior temporal sulcus (STS) has been reported to closely resemble the perceptual input, suggesting a rather part-based face representation as well. During a fMRI-experiment, we presented whole faces and facial parts of happy and fearful expressions. We used a new approach based on pattern similarity to analyze our data that has not been applied to this question before. We measured the similarity of activity patterns in OFA, FFA and STS between whole faces and facial parts, across and within emotions. We find part-based representation of facial expressions as well as emotion-indepented representations of whole faces in the FFA, which is in line with previous findings for identity recognition. The STS on the other hand exhibits an emotion-dependent representation of faces and facial parts. The OFA shows no specific effects but rather represents eye- and mouth-regions of both expressions similarly, supporting its role as an entry-point of facial information into the core network of face processing.

2.2. Introduction

There is quite some agreement that faces are processed somewhat differently from other objects as whole faces seem to be more than just the sum of their parts. This effect has been termed holistic processing, which is mostly described as the perception of a face as one ‘Gestalt’ that is not very much decomposed into its parts (Young et al., 1987; Tanaka and Farah, 1993; Farah et al., 1998, but see Richler et al., 2012). A long line of behavioral studies has investigated, how featural information - the shape, colour, size of facial parts -

as well as the configural information - the arrangement and metric distances of facial parts - are processed and integrated into a holistic percept of the face. Those findings indicate differences on how expressional and indentity-related information are processed. While information of the identity of a face is primarly represented holistically (Schiltz and Rossion, 2006; Schiltz et al., 2010; Rotshtein et al., 2007; Liu et al., 2010; Golarai et al., 2015; Soria Bauser and Suchan, 2015), expressional information rather seems to be decoded in a part-based fashion (McKelvie, 1995; Prkachin, 2003; Calvo and Nummenmaa, 2008; Derntl et al., 2009; Narme et al., 2011; Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011; Gosselin and Schyns, 2001; Blais et al., 2012).

The majority of studies targeting holistic face representation in the brain are however based on identity matching tasks. The FFA has been shown to respond both to whole face representations and configural information, as well as featural aspects (Rossion et al., 2000; Yovel and Kanwisher, 2004; Rotshtein et al., 2007; Schiltz and Rossion, 2006; Schiltz et al., 2010; Liu et al., 2010; Golarai et al., 2015). The OFA seems to contain mainly part-based (Liu et al., 2010; Golarai et al., 2015) and, only to a lesser extent, holistic (Schiltz and Rossion, 2006) representations. Based on behavioral findings and the proposal of distinct processing routes for identity and expression processing, it is likely that expressional information is represented differentially in the brain. However, this issue has hardly been investigated yet. Calvo and Beltrán (2014) detected modulations of face-specific right-hemispheric event-related potentials (ERP) by expressional whole faces, while left-sided ERPs were rather influenced by their parts. The timing of the ERPs though depended on the facial part and expression, with happy eyes eliciting earlier activation than angry eyes or other parts. Using an fMRI-adaptation paradigm with a composite face task of various expressions, Flack et al., (2015) found that the STS rather reflected a part-based representation of facial expressions. In contrast, the FFA adapted stronger to aligned faces indicating holistic representation, while the OFA did not show any differences between conditions.

In this study, we investigated the neuronal representation of expressional faces and their corresponding facial parts. We target this question using a novel approach based on pattern similarity which hasn´t applied to this question before. We based our approach on two previous studies investigating the neuronal representation of the integration of object parts into whole objects. MacEvoy and Epstein (2011) predicted neural patterns evoked by different classes of scenes (i.e. kitchen versus street scenes) from the neural patterns of their constituting objects (such as fridges, cars, traffic lights) and showed that using an average pattern of two objects leads to higher decoding accuracy than using the pattern of just one object. Kaiser et al., (2014) adopted this paradigm to successfully model whole-person evoked fMRI-patterns by a linear combination of neural patterns evoked by faces and headless bodies in the fusiform gyrus.

Here we adapted the design of MacEvoy and Epstein (2011) to study the integration of facial parts into whole faces. In particular, we used happy and fearful whole faces and stimuli of facial parts from these emotional faces showing either the eyes- or the mouth-region. We tested if patterns of full faces could be predicted by the average pattern of their emotional congruent parts. Additionally we investigated if an area has a stronger full face, i.e. potentially holistic preference which is independent of the presented emotion or, if an area stronger represents the expression of a face, irrespectively if a full face or just facials part were shown. As there is also evidence that eyes and mouth are differentially important for the recognition of different facial expressions (Calvo and Nummenmaa, 2008; Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011), we further tested, whether the eyes-region or the mouth-region were more similar to a whole faces.

2.3 Methods

Subjects

20 healthy subjects with normal or corrected-to-normal vision participated in this study. Three subjects were excluded due to excessive head-movement. Therefore, data of 17 subjects (mean age 30 ± 8 years, 8 males) entered the final analyses. The study was conducted according to the declaration of Helsinki, and was approved by the local ethics committee of the University of Tübingen. Subjects provided written consent prior to participation.

MRI data acquisition

Image acquisition of the first 15 subjects was performed on a Siemens 3T TIM Trio system (Siemens, Erlangen, Germany) with a 12-channel phased-array head coil. Due to a scanner-failure, data for the final 5 subjects were acquired following an upgrade to a Siemens Magnetom PRISMA 3T system (Siemens, Erlangen, Germany) using a 64-channel phased-array head coil. Since all analyses here are based on similarity of voxel patterns the difference in acquisition should not influence our results. T2*-weighted images from all subjects were collected using the same echo-planar imaging (EPI) sequence on both scanner types with TR = 2.3 s, TE = 35 ms, a flip angle of 79° and 32 slices with a voxel size of 3 × 3 × 3 mm. In addition, T1-weighted high-resolution anatomical images were obtained with a resolution of 1 × 1 × 1 mm.

Stimuli and experimental procedure

Stimuli were back-projected on a screen of 24° - 18° visual degrees, viewed via a tilted mirror. All stimuli were presented using Cogent Graphics 1.30 developed by John Romaya at the Wellcome Department of Imaging Neuroscience (http://vislab.ucl.ac.uk/cogent.php) running on MATLAB 2007a on a Windows PC.

The main experiment consisted of eight conditions: happy, neutral and fearful whole faces, happy and fearful eyes, happy and fearful mouth as well as scrambled pictures. Stimuli were taken from the NimStim-Face-Stimulus-Set (Tottenham et. al., 2002) and included faces of eleven actors. All pictures were presented gray-scale on top of a isoluminant and iso-contrast fourier scrambled background. The means of luminance and of its spatial variance, i.e. root-mean-square (RMS) contrast) were equaled across all whole face stimuli (luminance 126 cd/m2, RMS contrast 75,6 cd/m2) before cutting out the parts. The size of the stimulus was adjusted so that the subjects were able to grasp the stimulus at one glance without eye-movements to avoid a processing bias towards one facial parts. Full faces covered 5x4° visual degree and facial parts covered 2x4° visual degree. According to (Schultz et al., 2005), stimuli not greater than 3-5° of visual angle can be fully viewed (foveated) without eye movements. Additionally, phase-scrambled versions of full face stimuli were included as scrambled control condition. All stimuli, full faces, facial parts and scrambled pictures were presented centrally. Before the experiment, subjects familiarized with the stimuli in a short training session.

Paradigm

Conditions were presented in mini-blocks. During each mini-block six pictures of one category were presented from randomly drawn actors. Pictures were shown for 400 ms followed by a 600 ms blank with a fixation cross. In addition to the small stimulus size, the short duration of stimulus presentation was chosen to prevent eye-movements. After each mini-block subjects were asked to indicate which emotion had been shown. A screen, assigning each button to one of the three emotions appeared for a fixed window of 5000 ms. The sequence of one trial is also illustrated in Figure 1. Before the experiment, subjects were informed that eye and mouth conditions never included neutral faces. For scrambled conditions, subjects were told to press a random button to equalize the motor response. Each condition was repeated eight times in a pseudorandom, history-matched sequence resulting in 64 mini-blocks for each session. Each session started with a dummy mini-block that was discarded later but included to balance the condition sequence and to adapt initial neural responses. Each session lasted 12.3 minutes and subjects performed 6 of them. For technical reasons, two subjects were scanned with only 5 sessions.

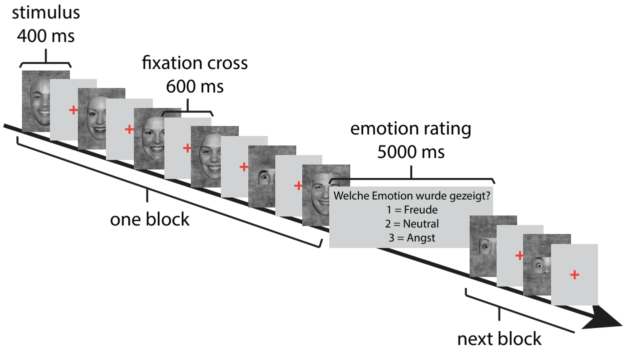


Fig. 1: Illustration of trial sequence. One trial consistet of a block of six stimuli of the same condition (duration of 400ms), interspaced by 600ms blank. After each blank, subjects had to press one of three buttons indication the emotion that was presented in the block.

Localizer experiment

Regions of interest (ROIs) related to visual face processing (occipital face area (OFA), fusiform face areas (FFA), superior temporal sulcus (STS)) and an additional ROI for early visual cortex (EVC) were defined using an independent functional block-design localizer with four conditions: static pictures of faces, of houses, of their static fourier scrambles (of both houses and faces) and movies of dynamic faces (soundless recordings of speech). All pictures and videos were presented grey-scale on top of a fourier scrambled background, and luminance and RMS contrast were equaled for all stimuli. The four conditions were presented in a block design, each block consisting of 5 stimulus presentations lasting 2400 ms, with inter-stimulus-intervals (ISIs) of 600ms. Blocks of each condition were repeated 8 times in a randomized, history-matched sequence. To ensure vigilance and balanced attention, subjects performed a one-back matching task to detect randomly occurring image repetitions that occurred with a frequency of 20%. The localizer comprised one session with a duration of about 8 minutes.

fMRI data preprocessing

Data were preprocessed using SPM (www.fil.ion.ucl.ac.uk/spm/). The first four images of each scanning session were discarded as dummy volumes to allow for equilibration of the T1 signal. Data were slice-time corrected for the differences in slice acquisition time and realigned to the first image to compensate for head motion. The structural image was coregistered to the mean functional image without normalization. Functional images were smoothed using a Gaussian Kernel with 3 mm full-width at half-maximum (FWHM). A high-pass filter of 128 s cut-off was applied to remove low-frequency signal drifts. In a last step, GLMdenoise (http://kendrickkay.net/GLMdenoise/) was applied to improve the signal-to-noise ratio of the data.

ROI definition

ROIs were defined using the independent localizer experiment for each subject individually. OFA and FFA were defined using the contrast “static face > house“, and STS was defined using the contrast “dynamic face > house“. If a ROI could not be defined using this method, we chose the contrast “static or dynamic faces > scrambled” instead. This approach has been applied before (Fox et al., 2009; Gschwind et al., 2012; Schultz et al., 2013). The latter definition may potentially result in fewer face-specific voxels (Schultz et al., 2013), but nevertheless allows for identification of face-responsive ROIs, for our purposes therefore a conservative ROI definition. To keep ROIs similar in size across participants, we used variable thresholding from p = 0.001 (uncorrected) to p= 0.05 (FWE corrected) (Fox et al., 2009; Jiang et al., 2006; Murray and Wojciulik, 2004). Additionally, a ROI of early visual cortex (EVC) was defined using voxels within a sphere of 15 mm radius around the voxel with the highest activation in the calcarine sulcus for the contrast “scramble > all face“. We were able to define the ROIs in the following number of subjects: left OFA in 17 subjects, right OFA in 16 subjects, left STS in 17 subjects, right STS in 15 subjects, left FFA in 10 subjects, right FFA in 12 subjects and EVC in 16 subjects.

Statistical analysis

Data of the localizer and the main experiment were analyzed in SPM5 using the general linear model (GLM) approach. The design matrix contained one regressor for each condition, modeled using a boxcar convolved with the canonical hemodynamic response function (HRF). Additionally, six realignment parameters obtained from the motion correction were included. For the main experiment we further added two regressors, one for the rating period and one for the button press of the subject. Each session of the main experiment was analyzed individually.

Pattern correlation analyses

For each subject and ROI, beta estimates of the main experiment were extracted separately from each session for each of the conditions as multivoxel patterns. This resulted in one pattern per session for each condition. Each pattern was Z-normalized by first subtracting its mean from each voxel and then dividing each voxel by the standard deviation of the pattern. As this normalization was done separately for each condition and each session, it did not induce any dependencies between conditions and sessions which could affect later correlation calculations. The patterns of each subject were then grouped into halves (each half including 3 sessions) and patterns within each half were averaged. This was done with every possible half- and half-data split (20 possible splits altogether). In order to test particular experimental hypotheses (see results section), we followed the procedures introduced by MacEvoy and Epstein (2011). In brief, they calculated euclidean distances (ED) between voxel patterns of a given ROI across two conditions (i.e. condition A and B) and then performed a classification analysis that determined the frequency (termed classification accuracy) with which ED of this condition pair was lower than the ED of another condition pair (i.e. condition A and C). They repeated the same analysis with correlations instead of EDs and obtained the same results. We therefore used correlations in our analysis. In line with their approach, each of the two patterns entering a correlation were obtained from two distinct halves of the data and correlations were calculated for all possible half-and-half data splits. Note that some patterns were average patterns of two conditions, i.e. the patterns of the eyes- and mouth-region were sometimes averaged to common pattern of facial parts. Correlations were Fisher-transformed before entering further analysis. Comparisons of the absolute values of this correlation coefficients were then done individually for each subject. For each analysis described in detail in the results section, the number of correct classifications (defined analogues to MacEvoy and Epstein, 2011) were summed up for every subject and divided by the number of all possible half-and-half data splits. This resulted in one accuracy percentage value for each subject. Accuracy values were then tested against 50% chance using one-sided, one sample t-tests unless stated otherwise. Results were Bonferroni-Holm-corrected for the number of ROIs.

2.4 Results

In the next section, we describe the detailed classification tests we used for our diffferent experimental questions. The accuracy in each ROI is given as percentage value that indicates the frequency of correct classification for each comparison.

Representation of Expression

First, we asked the very basic question if we could classify the emotional expressions from faces, the eyes- or the mouth-region. Classification procedure was performed by comparing the correlations of faces, eyes or mouth of the same emotion (i.e. happy eyes with happy eyes) with the correlations of faces, eyes or mouth of different emotions (i.e. happy eyes with fearful eyes), respectively (see Figure 2A-C). If the within-emotion correlation was higher than the between-emotion correlation, classification was considered as correct.

Results were averaged across emotions. Accuracy is given as percentage value that indicates the frequency of correct classification for each comparison.

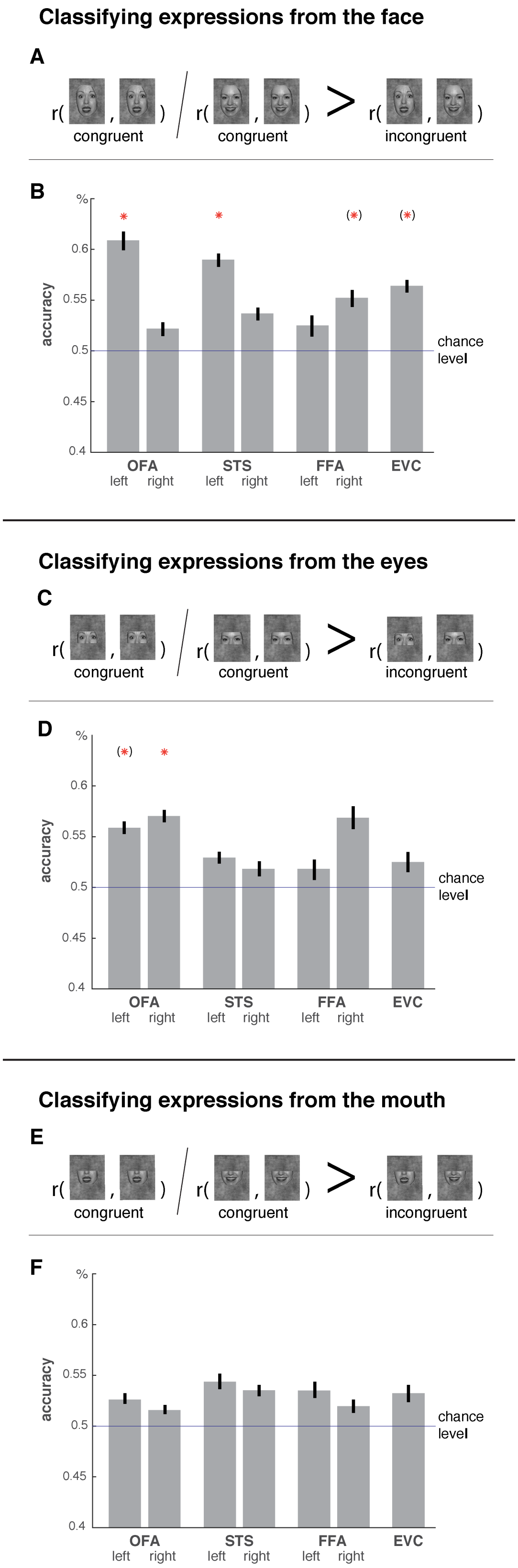


Fig. 2 A-F: Illustration of the classification procedure and results of expression representation. (A) Illustration of the expression classification from the face. (B) Results of the expression classification from the face. (C) Illustration of the expression classification from the eyes. (D) Results of the expression classification from the eyes. (E) Illustration of the expression classification from the mouth. (F) Results of the expression classification from the mouth. Results were always averaged across emotions. Error bars represent standard errors. *: $p < 0.05$ (one-sided, one sample t-tests against 50% chance), Bonferroni–Holm corrected for seven ROIs. (*): p–value did not survive correction.

Emotions could be classified correctly (see Figure 2D-F) from the whole face in the left OFA (accuracy: 61%, t(16) = 2.90, p = 0.005) and left STS (accuracy: 59%, t(16) = 3.37, p = 0.002) and from the eyes in the right OFA (accuracy: 57%, t(15) = 2.92, p = 0.005). The effects for correct classification of emotions from whole faces however did not survive Bonferoni-Holm-correction in the EVC (accuracy: 56 %, t(15) = 2.58, p = 0.01) and right FFA (accuracy: 55 %, t(11) = 1.80, p = 0.049) as well as from the eyes in the left OFA (accuracy: 56 %, t(16) = 2.33, p = 0.016).

'Constructing faces from facial parts'

Our second question was adopted from MacEvoy and Epstein (2011) who testet if patterns of a presented scene stimulus could be predicted by the average patterns of objects appearing in such scenes. Hence, they compared the euclidean distance of patterns of a scence (i.e. kitchen) and their constiuting objects (i.e. stove and refrigerator) to the ED between patterns of the same scence and objects from a different scene (i.e. car and traffic light). Analogue to their procedure we tested whether the voxel pattern obtained from a whole face of a given emotion (e.g. happy faces) correlated higher with the average pattern of face-parts belonging to the same emotion (e.g. happy eyes and happy mouth) than with face-parts of a different emotion (i.e. fearful eyes and fearful mouth), see Figure 3A for illustration. If this was the case, it would imply that the ROI encodes emotional features, not only from whole faces but also from its facial parts. Again, results were averaged across emotions.

We find a significant effect in the right STS (accuracy: 55%, t(14) = 2.89, p = 0.006). Further effects in left FFA (accuracy: 55%, t(9) = 2.74, p = 0.011) and right FFA (accuracy: 55%, t(11) = 2.26, p = 0.023) did not survive Bonferoni-Holm-correction (see Figure 3B).

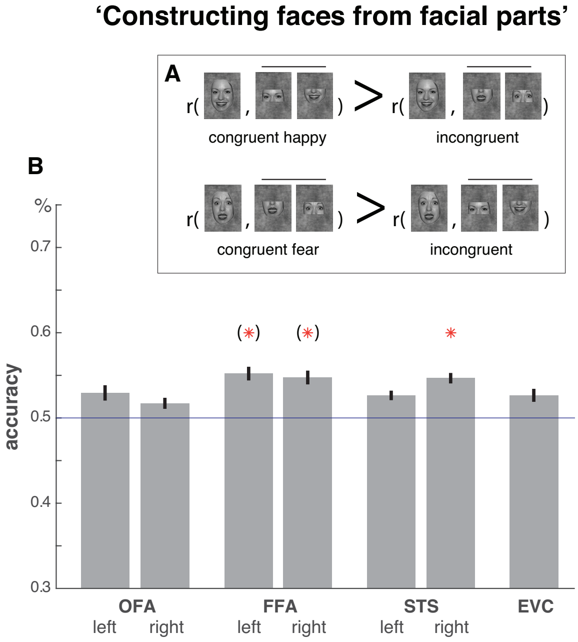


Fig. 3 Illustration and results of the classification procedure, adopted from MacEvoy and Epstein 2011 (A) Illustration of classification procedure. (B) Results of classification procedure, averaged across emotions. Error bars represent standard errors. *: $p < 0.05$ (one-sided, one sample t-tests against 50% chance), Bonferroni–Holm corrected for seven ROIs. (*): p-value did not survive correction.

Preference of emotion or visual input congruency

In the next step, we then wanted to test, if a an area stronger represents the expression, no matter if a whole faces or just facial parts are presented; or if the area has a stronger whole face, potentially holistic representation, that is emotion-independent. We compared the correlation of a face pattern of a given emotion (i.e. happy face) with the average patterns of emotional congruent face parts (i.e. happy eyes and happy mouth) to the correlation of the patterns of two whole faces with different emotions emotions (happy face with fearful face).

This comparison pinned congruency of emotion despite different visual input (e.g. happy face vs. happy parts) on the left side against congruency of visual input albeit with different emotions (happy full face with fearful fullface) on the right side. If the within face correlations (e.g. fear face - happy face) were higher than the correlation of faces and their emotional congruent parts (e.g. happy face - happy parts), it suggests that an area stronger represents whole faces even despite incongruent emotions (indicating a preference of congruent visual input, see Figure 4A). If the opposite pattern could be found, it would imply, that the presence of congruent emotions influenced patterns more than the visual input (indicating a preference of emotion congruency, see Figure 4B). Results were averaged across emotions.

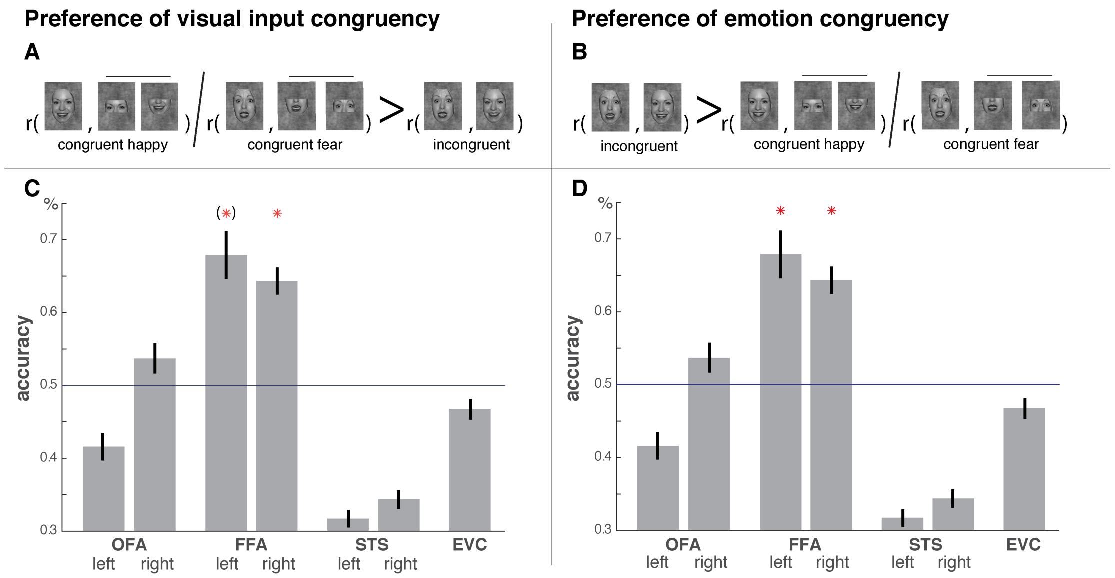


Fig. 4 A-D: Illustration of the classification procedure and results of the preference of emotion or visual input congruency. (A) Illustration of the preference of visual input congruency. (B) Illustration of the preference of emotion congruency. (C) Results of the preference of visual input congruency. (D) Results of the preference of emotion congruency. Results were always averaged across emotions. Error bars represent standard errors. *: $p < 0.05$ (one-sided, one sample t-tests against 50% chance), Bonferroni–Holm corrected for seven ROIs. (*): p –value did not survive correction.

For the first comparison (preference of congruent visual input) we obtain significant effect for the right FFA (accuracy: 55%, t(9) = 2.74, p = 0.011), the effect for the left FFA (accuracy: 55%, t(9) = 2.74, p = 0.011) did not survive Bonferoni-Holm-correction (see Figure 4C).

For the second comparison (preference of emotion congruency) we find a significant effect for the left STS (59%, t(16) = 2.73, p = 0.007), the effect for the right STS (accuracy: 57%, t(14) = 2.04, p = 0.03) did not survive Bonferoni-Holm-correction (see Figure 4D).

The impact of facial parts on facial expression processing

There is evidence that eyes and mouth are differentially important for the recognition of facial expressions. In general, the mouth seems to be sufficient for the recognition of happy faces (Calvo and Nummenmaa, 2008) while the identification of fear or sadness rather relies on the eyes (Eisenbarth and

Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011).

We aimed to test if such effects could also be found in the neuronal decoding of emotional faces. We first compared the correlation of a face pattern (i.e. happy face) with the average pattern of the emotionally congruent facial parts (i.e. happy eyes and happy mouth) to the correlation of the same face pattern with the pattern of only one emotionally congruent facial part (i.e. either happy eyes or happy mouth). If the correlation of the face with the average part-pattern is higher than the correlation of with a single facial part pattern, the part that is absent in the single-part condition (in this case happy mouth) adds significant information to the average part-pattern (see Figure 5A-B and Figure 6A-B). To test, if one part has a a greater impact on representing the emotion in general, we secondly testet, if the correlation of the face with this single part pattern was higher than the correlation of the face with the pattern of the other facial part. Since previous literature describes differential influence of mouth and eyes for the recognition for facial expressions, the analyses were carried out separately for each expression type.

Information content of fearful eyes (comparison: face-parts > face-mouth)

We find significant effects for left OFA (accuracy: 66%, t(16) = 3.54, p = 0.001). Additional effects for right OFA (accuracy: 68%, t(15) = 2.62, p = 0.009) as well as left and right FFA (accuracy: 70%, t(9) = 2.51, p = 0.017 and accuracy: 68%, t(11) = 2.26, p = 0.022, respectively) did not survive Bonferoni-Holm-correction (see Figure 5C).

Information content fearful mouth (comparison: face-parts > face-eyes)

We find significant effects for right OFA (accuracy: 72%, t(15) = 3.91, p < 0.000), left and right STS (accuracy: 69%, t(16) = 3.46, p = 0.002 and accuracy: 77%, t(16) = 4.30, p < 0.000, respectively) as well as EVC (accuracy: 78%, t(16) = 7.65, p < 0.000). The effect found in the left OFA (accuracy: 66%, t(16) = 2.33, p = 0.017) did not survive Bonferoni-Holm-correction (see Figure 5D).

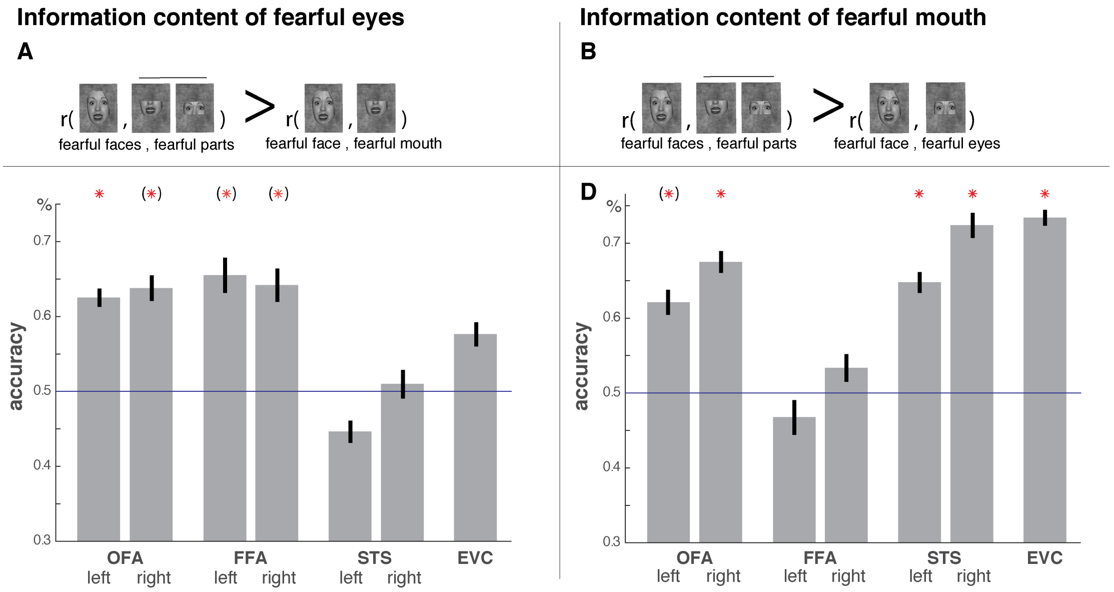


Fig. 5 A-D: Illustration of the classification procedure and results of the impact of facial parts on facial expression processing, fearful expressions. (A) Illustration of the information content of fearful eyes. (B) Illustration of the information content of fearful mouth. (C) Results of the information content of fearful eyes. (D) Results of the information content of fearful mouth. Error bars represent standard errors. *: p < 0.05 (one-sided, one sample t-tests against 50% chance), Bonferroni–Holm corrected for seven ROIs. (*): p –value did not survive correction.

Information content happy eyes (comparison: face-parts > face-mouth)

We find significant effects for right OFA (accuracy: 72%, t(15) = 4.73, p < 0.000), right STS (71%, t(16) = 4.05, p < 0.000) as well as left and right FFA (accuracy: 75%, t(9) = 4.48, p < 0.000 and accuracy: 74%, t(11) = 3.20, p = 0.004, respectively). The effect found in the left OFA (accuracy: 62%, t(16) = 1.85, p = 0.041) did not survive Bonferoni-Holm-correction (see Figure 6C).

Information content happy mouth (comparison: face-parts > face-eyes)

We find significant effects for left OFA (accuracy: 68%, t(16) = 2.87, p = 0.006) and EVC (accuracy: 71%, t(16) = 3.98, p < 0.000). The effects found for right OFA (accuracy: 63% t(15) = 2.14, p = 0.024) and left STS (accuracy: 63%, t(16) = 2.45, p = 0.013) did not survive Bonferoni-Holm-correction(see Figure 6C).

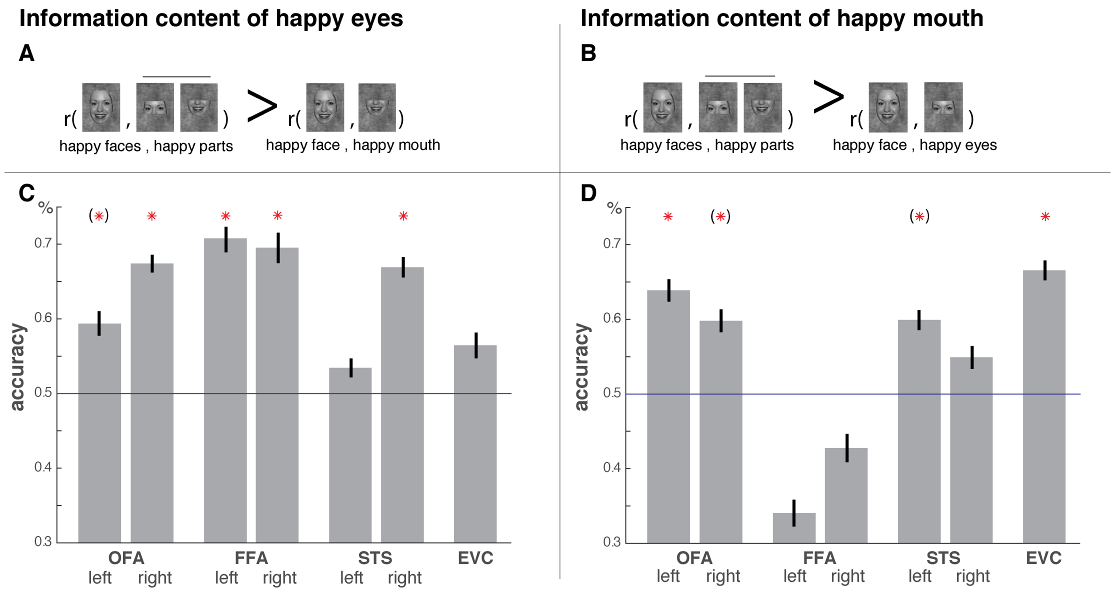


Fig. 6 A-D: Illustration of the classification procedure and results of the impact of facial parts on facial expression processing, happy expressions. (A) Illustration of the information content of happy eyes. (B) Illustration of the information content of happy mouth. (C) Results of the information content of happy eyes. (D) Results of the information content of happy mouth. Error bars represent standard errors. *: $p < 0.05$ (one-sided, one sample t-tests against 50% chance), Bonferroni–Holm corrected for seven ROIs. (*): p –value did not survive correction.

Stronger representation of eyes- or mouth region

The tests of the impact of the eyes- or mouth patterns on the correlation of the average part patterns with the whole face, do not give information if one facial part is indeed stronger represented in the whole face. Hence, we also tested the correlation of the whole face pattern with the eyes-pattern directly against the correlation of the full face pattern with the mouth-pattern. As we can infer the direction of the differences from the previous tests, we used one-sided paired-sampled ttests. Tests are again done seperately for both fearful and happy expressions (see Figure 6A for fearfull faces, 6B for happy faces). Table 1 contains the correlation values (mean ± standard error) of full faces pattern with eye pattern and mouth pattern for fearful expressions. Table 2 contains the correlation values (mean ± standard error) of full faces pattern with eye pattern and mouth pattern for fearful expressions.

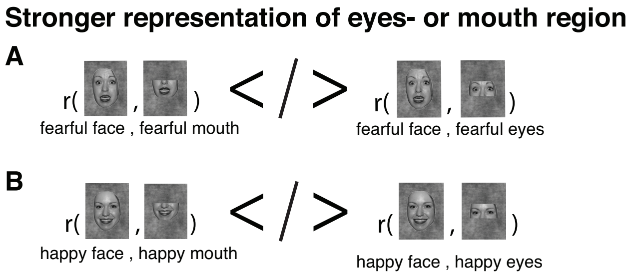


Fig. 6 A-B: Illustration of the comparison if the eyes- the or mouth regions are stronger represented in the full face. (A) fearfull faces. (B) happy faces

We find a significant effect in the left STS (t(16) = -2.8638, p = 0.0056) with higher correlations of the fearful mouth with the fearful face. The effect of the same direction found in the EVC (t(16) = -2.3629, p = 0.0160), did not survive Bonferoni-Holm-correction.

Table 1. Correlation values (mean ± standard error) of full faces pattern with eye pattern and mouth pattern for fearful expressions for all ROIs. Significant differences are marked in blue, differences that did not survive correction for multiple comparisons are marked in italic.

FEAR	**left OFA**	**right OFA**	**left STS**	**right STS**	**left FFA**	**right FFA**	***EVC***
full face with eyes	0.69 ± 0.02	0.91 ± 0.03	0.34 ± 0.02	0.58 ± 0.03	0.35 ± 0.01	0.53 ± 0.03	*0.60 ± 0.02*
full faces with mouth	0.68 ± 0.02	0.92 ± 0.03	0.39 ± 0.02	0.64 ± 0.03	0.31 ± 0.01	0.48 ± 0.02	*0.62 ± 0.02*

We find effects in the left and right FFA (happy: t(9) = 2.86, p = 0.0094, happy: t(11) = 2.24, p = 0.0235, respectively), showing higher correlations of happy eyes with the happy face, however non of them survived Bonferoni-Holm-correction.

Table 2. Correlation values (mean ± standard error) of full faces pattern with eye pattern and mouth pattern for happy expressions for all ROIs. Significant differences are marked in blue, differences that did not survive correction for multiple comparisons are marked in italic.

HAPPY	**left OFA**	**right OFA**	**left STS**	**right STS**	***left FFA***	***right FFA***	**EVC**
full face with eyes	0.65 ± 0.02	0.88 ± 0.03	0.35 ± 0.02	0.62 ± 0.04	*0.35 ± 0.01*	*0.52 ± 0.03*	0.63 ± 0.02
full faces with mouth	0.65 ± 0.02	0.87 ± 0.03	0.36 ± 0.02	0.60 ± 0.03	*0.25 ± 0.01*	*0.45 ± 0.03*	0.65 ± 0.02

2.5 Discussion

In our study we investigated how expressional faces and their corresponding facial parts are represented in the brain. In a first basic step, we tested if the two emotional expressions, fear and happiness, that we included in our experiment, could be classified from whole faces, the eyes or the mouth region (test of emotion classification). We then examined if the pattern of a whole face stimulus could be predicted by the average pattern of its emotional congruent parts. This question was adapted from the paper of MacEvoy and Epstein (2011) who predicted scene pattern from linearly combined pattern of their constituting objects ('Constructing faces from facial parts'). Afterwards, we analyzed if an area has a stronger full face, i.e. potentially holistic preference that generalizes across different emotional expressions of the faces or, if an area rather represents the expression, independently if a full face or just facial parts were shown (test of preference towards emotion- or visual input congruency). In a last step, we also investigated if the behaviorally described differential impact of eyes- and mouth-regions on the identifaction of different expressions (Calvo and Nummenmaa, 2008; Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011) is also reflected in the neuronal patterns of expressional faces and facial parts (Representation of mouth and eye regions). To our knowledge, this is the first study to investigate whole face as well as part-based representations of facial expressions in all three core areas of the face processing network using a pattern similarity approach.

Emotion classification

We were able to classify our two expressions happy and fear from static full faces in the OFA and STS, with stronger effects in the left hemisphere. The same effect in the FFA however, did not survive correction for multiple comparisons which might be due to the small number of subjects where we could reliably define the FFA from the localizer experiment.

In 2010, Said and colleagues (Said et al., 2010) sucessfully classified seven different dynamic facial expressions in posterior and anterior STS using bilateral ROIs. Unfortunately, the did not investigate OFA and FFA as their high-resolution-fMRI did not cover these brain parts. When Zhang et al., (2016) aimed to classify four basic facial expressions (neutral, fearful, angry, and happy) in face-responsive right-hemispheric brain areas, the STS only discriminated neutral from emotional but not within emotional expressions. The amygdala discriminated fearful from non-fearful faces and no effects were found for FFA or OFA. In contrast to our study, they used an event-related fmri-design and a face-irrelevant color-task. In our experiment we used a mini-block design and an explicit emotion recognition task, therefore enhancing the statistical power of the fmri-experiment and obtaining undisturbed attention towards the presented emotion. This might explain the differences between our study and the study of Zhang et al., (2016). The use of static and not dynamic face stimuli though might further explain our differences to the findings of Said et al., (2010). In contrast to both studies, we additionally presented isolated facial parts – namely eyes and mouth – which has not been investigated before. We were able to classify our two expression also from the eyes in the OFA.

'Constructing faces from facial parts'

One main aspect of our study was the question, if the pattern of a whole face could reliably be predicted by the average pattern of its emotional congruent parts. Analogues to the approach of MacEvoy and Epstein (2011), we aimed to predict a whole face pattern of a given emotion by the average pattern of its emotionally congruent parts. Prediction was successful in the right STS, the effects found for left and right FFA however, did not survive correction for multiple comparison. We consider this successful prediction as an indication, that the ROI encodes emotional information in both whole faces and facial parts.

Haxby et al., (2000) proposes a differential involvement of FFA and STS in the processing of identity and expressions and several studies have collected evidence for such a proposal (Furl et al., 2011; Hoffman and Haxby, 2000; Nestor et al., 2011; Puce et al., 1998; Said et al., 2010; Campbell et al., 2001). The distinction, however, appears less strict as previously assumed. The FFA has also been shown to represent emotional information as it responds stronger to the emotional than the neutral faces (Dolan et al., 1996; Gerber et al., 2008; Ishai et al., 2004; Pessoa et al., 2002; Surguladze et al., 2003; Vuilleumier and Pourtois, 2007; Vuilleumier et al., 2001; Winston et al., 2003; Atkinson et al., 2012). As both aspects, expression and identity, are always present in real faces, they have to be processed in parallel and generalized across each other. We usually do that effortlessly within an instant. Therefore, it is not surprising that the representation of this information somewhat overlaps on the neuronal level. The FFA further has been shown to contain both holistic as well as part-based facial representations (Rossion et al., 2000; Yovel and Kanwisher, 2004; Rotshtein et al., 2007; Schiltz and Rossion, 2006; Schiltz et al., 2010; Liu et al., 2010; Golarai et al., 2015) underlying its importance for processing faces (Kanwisher et al., 1997). Even though the effects we find for the FFA do not survive correction for multiple comparisions for the number of ROIs, they are in line with the described findings. As discussed before, the weaker effects in the FFA might be due to the small number of subjects, were the FFA could reliably defined from the localizer experiment.

The STS in contrast has been shown to contain rather part-based representation of facial expressions (Flack et al., 2015) and to be relatively unaffected by contrast reversal of faces, opposite to the FFA, which is affected by this manipulation (Harris et al., 2014). Behaviorally, contrast reversal impairs identity recognition but only slightly affects expression recognition (Bruce and Langton, 1994; Burton et al., 2005; Russell et al., 2006), adding evidence to potentially distinct holistic and part-based representation of facial identity and facial expression. Said et. al., (2010) further found the perceptual similarity of facial expression to be correlated to the similarity of their activity patterns in the STS. They therefore conclude that the STS closely resembles the perceptual input. Our results that the STS represents emotional information also in facial parts neatly fit into this findings.

Preference towards emotion- or visual input congruency

In a third step, we amined to detect preferences of the ROIs either towards the emotional content in our face stimuli, independently if whole faces or just facial parts were shown (preference of emotion congruency) or if the ROIs rather decodes whole faces, irrespectively of the expressions of the face (preference of congruent visual input). We find a preference towards emotion-independent whole face representation (preference of congruent visual input) in the FFA and the opposite pattern, a preference of emotion congruency in the STS.

This results extend our findings that the STS represents expressional information in facial parts by showing that it also responds stronger to the emotional content of a stimulus, independent if a full face or facial parts were shown. This adds further evidence for a strong part-based representation of facial information in the STS. In contrast, the findings for the FFA add evidence to a preferentially holistic face representation in the FFA (Rotshtein et al., 2007; Schiltz and Rossion, 2006; Schiltz et al., 2010; Soria Bauser and Suchan, 2015) which might be stronger involved in the representation of facial identity than facial expressions.

Representation of mouth and eye regions

Our last two tests aimned to confirm the behaviorally described differential inpact of eyes- and mouth-regions for the recognition of different expressions (Calvo and Nummenmaa, 2008; Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011).

Prior behavioural studies

Some behavioral studies have shown that happy faces can be easily recognized only by the mouth (Calvo and Nummenmaa, 2008) while the eyes seem to be more important for fear or sadness (Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011). There are, however, also contrasting studies. Blais et. al., (2012) found the mouth to be most important for recognition in both static and dynamic stimuli with no differences between emotions. Together, these effects are often taken as evidence for a mainly part-based representation of facial expressions. In contrast to this, studies using inversion paradigms (Derntl et al., 2009; McKelvie, 1995; Prkachin, 2003) or the composite face effect (Calder and Jansen, 2005; Calder et al., 2000; Tanaka et al., 2012) show that not only facial identity, but also facial expressions are affected by configural information and do not only rely on featural aspects. Additionally, in a free viewing emotion-labeling task, Guo (2012) found variations of fixation durations on facial parts depending on the emotion, however subjects still scanned the whole face before labeling the emotion. Similar effects were reported by Eisenbarth and Alpers (2011). Nevertheless, it remains unclear whether the scanned single features are indeed formed into a holistic percept of the face, or how the featural and configural information are actually integrated for recognition. Guo (2012) claims a holistic processing strategy, but as they used stimuli of varying intensities, they also offer the explanation that the investigation of several face parts might simply help to identify the correct expression, at least for ambiguous low intensity stimuli, as single facial parts might carry less diagnostic information when not at maximum level.

OFA and FFA

We find that in the OFA, both facial parts, eyes and mouth, add comparable information to the whole face representation and this effects seems to be emotion-independent. In contrast, in the FFA the eyes seem to have a greater informational content, especially for happiness.

Only a few studies investigated the representation of isolated facial parts in the core face network. The OFA has been shown to respond to facial parts but not being influenced by their configuration (Liu et al., 2010; Pitcher et al., 2007), and preferentially resembling the physical structure of faces (Rotshtein et al., 2005). This is in line with our results showing a rather equal information content of facial parts independent from the presented expression with no part being stronger represented than the other one. For the FFA however, we find that the eyes add more information to the average pattern of both facial parts than the mouth. The eyes have repeatedly been shown to be important for identity recognition (McKelvie, 1976; Gilad et al., 2009; Sormaz et al., 2013; Malcolm et al., 2008), albeit especially when presented within the whole face. Gilad et al., (2009) for example found recognition rates of about 90% for chimeric faces showing normal eyes, while the rest of the face was contrast-negated, in contrast to about 15% when only eyes were presented. The FFA displayed a similar response pattern with indistinguishable activation towards normal and chimeric faces, while activation for eyes-only and fully contrast negated faces was significantly smaller. Similar behavioral results were reported by Sormaz et. al., (2013). As the FFA has been described to be preferentially involved in identity processing (Furl et al., 2011; Hoffman and Haxby, 2000; Nestor et al., 2011), it might not be surprising that we find such an effect of the eye-region. Another possible explanation here could be feedback from amygdala though. The amygdala has been shown to have connections to the FFA (Smith et al., 2009), has been discussed as potential modulator of FFA activity (Vuilleumier et al., 2004) and to be involved the processing of emotional faces (Morris et al., 1998). Furthermore, eyes alone have been effective enough to elicit activation in the amygdala (Morris et al., 2002; Whalen et al., 2004). Since we did not optimize our fMRI protocol to obtain signal from the amygdala that is often affected by sensitivity losses due to susceptibility artifacts, we were unable to investigate amygdala activity.

STS

The effects found for the STS show a more complex picture. While for the left STS, the mouth has a greater informational value than the eyes independent from the expression type, the right STS seems to differentiate between the expressions. The mouth seem to have a higher information value for fearful expressions, while the eyes seem to have a higher information value for happy expressions.

The majority of behavioral studies describe the eyes as highly diagnostic for fearful faces and the mouth as highly diagnostic for happy expressions (Calvo and Nummenmaa, 2008; Eisenbarth and Alpers, 2011; Adolphs et al., 2005; Smith et al., 2005; Gosselin et al., 2011). The right STS however shows the opposite pattern while in the left STS the mouth is generally stronger represented than the eyes. As mentioned above, Blais et al., (2012) found the mouth but not the eyes to be correlated most with the accuracy of expression recognition. Blais et al., (2012) further conducted a motion analysis showing that the mouth contains the highest variance between expressions, and hence the most information for expression discrimination. They propose that brain has developed some sort of strategy using the information from the mouth as we are usually exposed to moving and not static faces in real world. This strategy however is applied also when seeing static expressions in experimental setups. Another interesting study supporting the impact of the mouth for expression recognition was conducted by Malcolm et al., (2008) who measured fixation time during a face recognition task. Subjects had to choose the more similar face out of a pair of probe faces, with none of them being identical to the target face. Subjects had to base their response either on identity similarity or expression similarity. If decisions were based on identity similarity, subjects focused stronger on the upper half of the face and when their decisions were based on expression similarity, they focused stronger on the lower half of the face. Furthermore, the STS has been shown to be activated by different types of mouth movement like linguistic but also non-linguistic movements (Campbell and Capek, 2008; Puce et al., 2003). It is therefore not surprising that it contains a strong mouth representation (see also Andrews and Ewbank, 2004).

Limitations

One limitation of our experimental design is that comparisons between whole faces and facial parts might be confounded by low level aspects of the images. Facial part stimuli are smaller in size than the full face stimuli. On the other hand, differences in the expressional information could be treated as high-level confounds. Our third test of a preference of emotional or visual input congruency explicitly pinnes these low-level aspects of stimulus size while keeping the expressional information constant against variations of emotional expressions while keeping the stimulus size identical. We find differential effects for these tests for FFA and STS that generalize across one or the other potential confound. Previous studies comparing whole faces with facial parts were also facing the same problem of low-level confounds of stimulus size (Anzellotti and Caramazza, 2015; Arcurio et al., 2012). Anzellotti and Caramazza (2015) presented full faces as well as upper, lower, left and right face halves during an identity recognition task. In a similar approach to ours, he reconstructed whole face pattern from linear combinations of the activity pattern of facial halves. Like us, they included the early visual cortex as control areas in their analysis. They successfully modelled whole faces from left and right face halves in V1, even with the corresponding hemispheric bias. We also include a control ROI of the EVC in our analysis. We only find a stronger representation of the mouth for both expression types. However, this does not have to be related to low-level confounds of the stimulu. This effects in EVC are similar to the effects found in the STS, which also shows a strong mouth-preference. Early visual areas have been shown to be affected by high-level feedback (Datta and DeYoe, 2009; Fischer and Whitney, 2009; Ress et al., 2000; Serences, 2008; Fang et al., 2008; Murray et al., 2002), even from face-processing areas (Hsieh et al., 2010; Strother et al., 2011; Ayzenshtat et al., 2012). Strother et al., (2011) found longer visual persistence for upright but not inverted face stimuli not only in face but also in the retinotopic visual areas. Petro et al., (2013) presented whole faces with variations in eyes and mouth and asked subjects to detect the resulting changes of gender and expression. They found not only regions in V1 to be task-modulated that represented the changed parts, but surprisingly, also the remaining division of V1 contained task-specific information. A recent study by Ayzenshtat et al., (2012) reported correlated activity in neural populations of monkey V1 with the behavior of the animals using monkey face stimuli. As this correlation between V1 activity and the perceptual processing of faces was found only for late but not early responses of V1, the effects is thought to be mediated by high-level feedback mechanism. In line with this, anatomical feedback projections have been shown from temporal cortices to V1, including projections from the temporal sulcus (Rockland and Van Hoesen, 1994).

2.6 Conclusion

Our study is the first to explicitly investigate part-based and whole face representations of facial expressions in the core network of face processing using a pattern similarity approach. We find part-based representation of facial expressions in the FFA. However, we also find a preference to whole faces that is emotion-independent which matches with previous findings on holistic face representation in the FFA (Yovel and Kanwisher, 2004; Rotshtein et al., 2007; Schiltz and Rossion, 2006; Schiltz et al., 2010; Soria Bauser and Suchan, 2015). In line with the importance of the eyes-region for identity recognition (McKelvie, 1976; Gilad et al., 2009; Sormaz et al., 2013; Malcolm et al., 2008) and the described involvement of the FFA in identity processing (Furl et al., 2011; Hoffman and Haxby, 2000; Nestor et al., 2011), we further find a trend of a stronger representation of the eyes-region. The STS, in contrast shows the previously proposed emotion-dependent representation of faces and facial parts. Our findings neatly fit into the existing literature and supports its role in expression recognition. The findings for OFA support its role as an entry point of facial information (Pitcher et al., 2007; Pitcher et al., 2011) into the core network. We neither detect specific effects for emotion-independent whole face representations like in the FFA, nor specific preferences for expressional information like in STS. Instead we find rather similar representation of eyes and mouth for both expressions, with non being more dominant than other.

2.7 References

Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P., Damasio, A.R., 2005. A mechanism for impaired fear recognition after amygdala damage. Nature 433, 68–72.

Andrews, T.J., Ewbank, M.P., 2004. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. NeuroImage 23, 905–913.

Anzellotti, S., Caramazza, A., 2015. From Parts to Identity: Invariance and Sensitivity of Face Representations to Different Face Halves. Cerebral Cortex bhu337.

Arcurio, L.R., Gold, J.M., James, T.W., 2012. The response of face-selective cortex with single face parts and part combinations. Neuropsychologia 50, 2454–2459.

Atkinson, A.P., Vuong, Q.C., Smithson, H.E., 2012. Modulation of the face- and body-selective visual regions by the motion and emotion of point-light face and body stimuli. NeuroImage 59, 1700–1712.

Ayzenshtat, I., Gilad, A., Zurawel, G., Slovin, H., 2012. Population response to natural images in the primary visual cortex encodes local stimulus attributes and perceptual processing. Journal of Neuroscience 32, 13971–13986.

Blais, C., Roy, C., Fiset, D., Arguin, M., Gosselin, F., 2012. The eyes are not the window to basic emotions. Neuropsychologia 50, 2830–2838.

Bruce, V., Langton, S., 1994. The use of pigmentation and shading information in recognising the sex and identities of faces. Perception 23, 803–822.

Burton, A.M., Jenkins, R., Hancock, P.J.B., White, D., 2005. Robust representations for face recognition: the power of averages. Cogn Psychol 51, 256–284.

Calder, A.J., Jansen, J., 2005. Configural coding of facial expressions: The impact of inversion and photographic negative. Visual Cognition.

Calder, A.J., Young, A.W., Keane, J., Dean, M., 2000. Configural information in facial expression perception. J Exp Psychol Hum Percept Perform 26, 527–551.

Calvo, M.G., Beltrán, D., 2014. Brain lateralization of holistic versus analytic processing of emotional facial expressions. NeuroImage 92, 237–247.

Calvo, M.G., Nummenmaa, L., 2008. Detection of emotional faces: salient physical features guide effective visual search. J Exp Psychol Gen 137, 471–494.

Campbell, R., Capek, C., 2008. Seeing speech and seeing sign: insights from a fMRI study. Int J Audiol 47 Suppl 2, S3–9.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M.J., David, A.S., 2001. Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). Brain research Cognitive brain research 12, 233–243.

Datta, R., DeYoe, E.A., 2009. I know where you are secretly attending! The topography of human visual attention revealed with fMRI. Vision Res 49, 1037–1044.

Derntl, B., Habel, U., Windischberger, C., Robinson, S., Kryspin-Exner, I., Gur, R.C., Moser, E., 2009. General and specific responsiveness of the amygdala during explicit emotion recognition in females and males. BMC Neurosci 10, 91–14.

Dolan, R.J., Fletcher, P., Morris, J., Kapur, N., Deakin, J.F., Frith, C.D., 1996. Neural activation during covert processing of positive emotional facial expressions. NeuroImage 4, 194–200.

Eisenbarth, H., Alpers, G.W., 2011. Happy mouth and sad eyes: scanning emotional facial expressions. Emotion 11, 860–865.

Fang, F., Kersten, D., Murray, S.O., 2008. Perceptual grouping and inverse fMRI activity patterns in human visual cortex. JOV 8, 2.1–9.

Farah, M.J., Wilson, K.D., Drain, M., Tanaka, J.N., 1998. What is "special" about face perception? Psychol Rev 105, 482–498.

Fischer, J., Whitney, D., 2009. Attention narrows position tuning of population responses in V1. Curr Biol 19, 1356–1361.

Flack, T.R., Andrews, T.J., Hymers, M., Al-Mosaiwi, M., Marsden, S.P., Strachan, J.W.A., Trakulpipat, C., Wang, L., Wu, T., Young, A.W., 2015. Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. Cortex 69, 14–23.

Fox, C.J., Iaria, G., Barton, J.J.S., 2009. Defining the face processing network: Optimization of the functional localizer in fMRI. Hum Brain Mapp 30, 1637–1651.

Furl, N., Garrido, L., Dolan, R.J., Driver, J., Duchaine, B., 2011. Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. J Cogn Neurosci 23, 1723–1740.

Gerber, A.J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J., Peterson, B.S., 2008. An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. Neuropsychologia 46, 2129–2139.

Gilad, S., Meng, M., Sinha, P., 2009. Role of ordinal contrast relationships in face encoding. Proceedings of the National Academy of Sciences 106, 5353–5358.

Golarai, G., Ghahremani, D.G., Eberhardt, J.L., Gabrieli, J.D.E., 2015. Distinct representations of configural and part information across multiple face-selective regions of the human brain. Front. Psychology 6, 1710.

Gosselin, F., Schyns, P.G., 2001. Bubbles: a technique to reveal the use of information in recognition tasks. Vision Res 41, 2261–2271.

Gosselin, F., Spezio, M.L., Tranel, D., Adolphs, R., 2011. Asymmetrical use of eye information from faces following unilateral amygdala damage. Social Cognitive and

Affective Neuroscience 6, 330–337.

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., Vuilleumier, P., 2012. White-matter connectivity between face-responsive regions in the human brain. Cerebral Cortex 22, 1564–1576.

Guo, K., 2012. Holistic gaze strategy to categorize facial expression of varying intensities. PLoS ONE 7, e42585.

Harris, R.J., Young, A.W., Andrews, T.J., 2014. Brain regions involved in processing facial identity and expression are differentially selective for surface and edge information. NeuroImage 97, 217–223.

Haxby, J., Hoffman, E., Gobbini, M., 2000. The distributed human neural system for face perception. Trends Cogn Sci (Regul Ed) 4, 223–233.

Hoffman, E.A., Haxby, J.V., 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nat Neurosci 3, 80–84.

Hsieh, P.-J., Vul, E., Kanwisher, N., 2010. Recognition alters the spatial pattern of FMRI activation in early retinotopic cortex. Journal of Neurophysiology 103, 1501–1507.

Ishai, A., Pessoa, L., Bikle, P.C., Ungerleider, L.G., 2004. Repetition suppression of faces is modulated by emotion. Proc Natl Acad Sci USA 101, 9827–9832.

Jiang, X., Rosen, E., Zeffiro, T., Vanmeter, J., Blanz, V., Riesenhuber, M., 2006. Evaluation of a shape-based model of human face discrimination using FMRI and behavioral techniques. Neuron 50, 159–172.

Kaiser, D., Strnad, L., Seidl, K.N., Kastner, S., Peelen, M.V., 2014. Whole person-evoked fMRI activity patterns in human fusiform gyrus are accurately modeled by a linear combination of face- and body-evoked activity patterns. Journal of Neurophysiology 111, 82–90.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17, 4302–4311.

Liu, J., Harris, A., Kanwisher, N., 2010. Perception of face parts and face configurations: an FMRI study. J Cogn Neurosci 22, 203–211.

MacEvoy, S.P., Epstein, R.A., 2011. Constructing scenes from objects in human occipitotemporal cortex. Nat Neurosci 14, 1323–1329.

Malcolm, G.L., Lanyon, L.J., Fugard, A.J.B., Barton, J.J.S., 2008. Scan patterns during the processing of facial expression versus identity: an exploration of task-driven and stimulus-driven effects. JOV 8, 2.1–9.

McKelvie, S.J., 1995. Emotional expression in upside-down faces: evidence for configurational and componential processing. Br J Soc Psychol 34 (Pt 3), 325–334.

McKelvie, S.J., 1976. The role of eyes and mouth in the memory of a face. The American Journal of Psychology 89, 311–323.

Morris, J.S., deBonis, M., Dolan, R.J., 2002. Human amygdala responses to fearful eyes. NeuroImage 17, 214–222.

Morris, J.S., Friston, K.J., Büchel, C., Frith, C.D., Young, A.W., Calder, A.J., Dolan, R.J., 1998. A neuromodulatory role for the human amygdala in processing emotional facial expressions. Brain 121, 47–57.

Murray, S.O., Kersten, D., Olshausen, B.A., Schrater, P., Woods, D.L., 2002. Shape perception reduces activity in human primary visual cortex. Proc Natl Acad Sci USA 99, 15164–15169.

Murray, S.O., Wojciulik, E., 2004. Attention increases neural selectivity in the human lateral occipital complex. Nat Neurosci 7, 70–74.

Narme, P., Bonnet, A.-M., Dubois, B., Chaby, L., 2011. Understanding facial emotion perception in Parkinson's disease: the role of configural processing. Neuropsychologia 49, 3295–3302.

Nestor, A., Plaut, D.C., Behrmann, M., 2011. Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. Proceedings of the National Academy of Sciences 108, 9998–10003.

Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G., 2002. Neural processing of emotional faces requires attention. Proc Natl Acad Sci USA 99, 11458–11463.

Petro, L.S., Smith, F.W., Schyns, P.G., Muckli, L., 2013. Decoding face categories in diagnostic subregions of primary visual cortex. Eur J Neurosci 37, 1130–1139.

Pitcher, D., Walsh, V., Duchaine, B., 2011. The role of the occipital face area in the cortical face perception network. Exp Brain Res 209, 481–493.

Pitcher, D., Walsh, V., Yovel, G., Duchaine, B., 2007. TMS evidence for the involvement of the right occipital face area in early face processing. Curr Biol 17, 1568–1573.

Prkachin, G.C., 2003. The effects of orientation on detection and identification of facial expressions of emotion. Br J Psychol 94, 45–62.

Puce, A., Allison, T., Bentin, S., Gore, J.C., Mccarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. J Neurosci 18, 2188–2199.

Puce, A., Syngeniotis, A., Thompson, J.C., Abbott, D.F., Wheaton, K.J., Castiello, U., 2003. The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. NeuroImage 19, 861–869.

Ress, D., Backus, B.T., Heeger, D.J., 2000. Activity in primary visual cortex predicts performance in a visual detection task. Nat Neurosci 3, 940–945.

Richler, J.J., Palmeri, T.J., Gauthier, I., 2012. Meanings, mechanisms, and measures of holistic processing. Front. Psychology 3, 553.

Rockland, K.S., Van Hoesen, G.W., 1994. Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. Cereb Cortex 4, 300–313.

Rossion, B., Dricot, L., Devolder, A., Bodart, J.M., Crommelinck, M., de Gelder, B., Zoontjes, R., 2000. Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. J Cogn Neurosci 12, 793–802.

Rotshtein, P., Geng, J.J., Driver, J., Dolan, R.J., 2007. Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: behavioral and functional magnetic resonance imaging data. J Cogn Neurosci 19, 1435–1452.

Rotshtein, P., Henson, R.N.A., Treves, A., Driver, J., Dolan, R.J., 2005. Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. Nat Neurosci 8, 107–113.

Russell, R., Sinha, P., Biederman, I., Nederhouser, M., 2006. Is pigmentation important for face recognition? Evidence from contrast negation. Perception.

Said, C.P., Moore, C.D., Engell, A.D., Todorov, A., Haxby, J.V., 2010a. Distributed representations of dynamic facial expressions in the superior temporal sulcus. JOV 10, 11–11.

Schiltz, C., Dricot, L., Goebel, R., Rossion, B., 2010. Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. JOV 10, 25.1–16.

Schiltz, C., Rossion, B., 2006. Faces are represented holistically in the human occipito-temporal cortex. NeuroImage 32, 1385–1394.

Schultz, J., Brockhaus, M., Bulthoff, H.H., Pilz, K.S., 2013. What the Human Brain Likes About Facial Motion. Cereb Cortex 23, 1167–1178.

Schultz, R.T., 2005. Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. International Journal of Developmental Neuroscience 23, 125–141.

Serences, J.T., 2008. Value-based modulations in human visual cortex. Neuron 60, 1169–1181.

Smith, C.D., Lori, N.F., Akbudak, E., Sorar, E., Gultepe, E., Shimony, J.S., McKinstry, R.C., Conturo, T.E., 2009. MRI diffusion tensor tracking of a new amygdalo-fusiform and hippocampo-fusiform pathway system in humans. J Magn Reson Imaging 29, 1248–1261.

Smith, M.L., Cottrell, G.W., Gosselin, F., Schyns, P.G., 2005. Transmitting and decoding facial expressions. Psychological science : a journal of the American Psychological Society / APS 16, 184–189.

Soria Bauser, D., Suchan, B., 2015. Is the whole the sum of its parts? Configural processing of headless bodies in the right fusiform gyrus. Behav Brain Res 281, 102–110.

Sormaz, M., Andrews, T.J., Young, A.W., 2013. Contrast negation and the importance of the eye region for holistic representations of facial identity. J Exp Psychol Hum Percept Perform 39, 1667–1677.

Strother, L., Mathuranath, P.S., Aldcroft, A., Lavell, C., Goodale, M.A., Vilis, T., 2011. Face inversion reduces the persistence of global form and its neural correlates. PLoS ONE 6, e18705.

Surguladze, S.A., Brammer, M.J., Young, A.W., Andrew, C., Travis, M.J., Williams, S.C.R., Phillips, M.L., 2003. A preferential increase in the extrastriate response to signals of danger. NeuroImage 19, 1317–1328.

Tanaka, J.W., Farah, M.J., 1993. Parts and wholes in face recognition. Q J Exp Psychol A 46, 225–245.

Tanaka, J.W., Kaiser, M.D., Butler, S., Le Grand, R., 2012. Mixed emotions: holistic and analytic perception of facial expressions., in:. Presented at the Cognition & Emotion, pp. 961–977.

Tottenham, N., Tanaka, J.W., Leon, A.C., McCarry, T., Nurse, M., Hare, T.A., Marcus, D.J. Westerlund, A., Casey, B.J. Nelson, C., 2009. The NimStim set of facial expressions: Judgments from untrained research participants. Psychiatry Research 168, 242–249

Vuilleumier, P., Armony, J.L., Driver, J., Dolan, R.J., 2001. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30, 829–841.

Vuilleumier, P., Pourtois, G., 2007. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. Neuropsychologia 45, 174–194.

Vuilleumier, P., Richardson, M.P., Armony, J.L., Driver, J., Dolan, R.J., 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. Nat Neurosci 7, 1271–1278.

Whalen, P.J., Kagan, J., Cook, R.G., Davis, F.C., Kim, H., Polis, S., McLaren, D.G., Somerville, L.H., McLean, A.A., Maxwell, J.S., Johnstone, T., 2004. Human amygdala responsivity to masked fearful eye whites. Science 306, 2061–2061.

Winston, J.S., O'Doherty, J., Dolan, R.J., 2003. Common and distinct neural responses during direct and incidental processing of multiple facial emotions. NeuroImage 20, 84–97.

Young, A.W., Hellawell, D., Hay, D.C., 1987. Configurational information in face perception. Perception 16, 747–759.

Yovel, G., Kanwisher, N., 2004. Face perception: domain specific, not process specific. Neuron 44, 889–898.

Zhang, H., Japee, S., Nolan, R., Chu, C., Liu, N., Ungerleider, L.G., 2016. Face-selective regions differ in their ability to classify facial expressions. NeuroImage 130, 77–90.

3 Face processing regions are sensitive to distinct aspects of temporal sequence in facial dynamics

This chapter has been reproduced from an article published in Neuroimage: Reinl, M., and Bartels, A. (2014). Face processing regions are sensitive to distinct aspects of temporal sequence in facial dynamics. Neuroimage 102, 407–415.

3.1 Abstract

Facial movement conveys important information for social interactions, yet its neural processing is poorly understood. Computational models propose that shape- and temporal sequence sensitive mechanisms interact in processing dynamic faces. While face processing regions are known to respond to facial movement, their sensitivity to particular temporal sequences has barely been studied. Here we used fMRI to examine the sensitivity of human face-processing regions to two aspects of directionality in facial movement trajectories. We presented genuine movie recordings of increasing and decreasing fear expressions, each of which were played in natural or reversed frame order. This two-by-two factorial design matched low-level visual properties, static content and motion energy within each factor, emotion-direction (increasing or decreasing emotion) and timeline (natural versus artificial). The results showed sensitivity for emotion-direction in FFA, which was timeline-dependent as it only occurred within the natural frame order, and sensitivity to timeline in the STS, which was emotion-direction-dependent as it only occurred for decreased fear. The occipital face area (OFA) was sensitive to the factor timeline. These findings reveal interacting temporal sequence sensitive mechanisms that are responsive to both ecological meaning and to prototypical unfolding of facial dynamics. These mechanisms are temporally directional, provide socially relevant information regarding emotional state or naturalness of behavior, and agree with predictions from modeling and predictive coding theory.

3.2 Introduction

Social interactions in real life are dynamic by nature, and numerous social signals, including those conveyed by faces, rely on timing and temporal sequences. In accord with this, evidence shows that temporal contingencies in facial dynamics play an important role in behavior. Facial dynamics can improve the recognition of subtle emotional expressions (Ambadar et al., 2005; Wehrle et al., 2000), the personal identity of others (O'Toole et al., 2002; Thornton and Kourtzi, 2002), gender (Hill and Johnston, 2001) and language (Campbell, 1992).

Despite their importance, dynamic face stimuli have not been studied nearly as extensively as their static counterparts. Brain regions responsive to static faces increase their activity in response to facial dynamics (Fox et al., 2009; Kilts et al., 2003; LaBar et al., 2003; Sato et al., 2004; Schultz and Pilz, 2009; Trautmann et al., 2009). Particularly the posterior superior temporal sulcus (pSTS) is more sensitive to facial dynamics than the fusiform face area (FFA) or the occipital face area (OFA) (Pitcher et al., 2011; Schultz et al., 2012). This partial regional differentiation parallels one proposed for the encoding of changeable versus non-changeable aspects of faces (Haxby et al., 2000; Ishai et al., 2005; Puce and Perrett, 2003; Said et al., 2011; Vuilleumier and Pourtois, 2007). Changeable aspects of a face have been shown to be encoded by the pSTS, such as emotional expression (Said et al., 2010), gaze-direction (Hoffman and Haxby, 2000; Puce et al., 1998), mouth movements (Campbell et al., 2001) and intention (Nummenmaa and Calder, 2009) while non-changeable aspects are thought to be mainly processed by the FFA, such as identity (Furl et al., 2011; Hoffman and Haxby, 2000; Nestor et al., 2011; Steeves et al., 2009) (but see Kriegeskorte et al., 2007), race (Natu et al., 2011), and gender (Kaul et al., 2011).

It is nevertheless unclear which aspects of facial dynamics drive the increasing responses in the core face processing regions. Schultz et al. (2012) showed that the amount of static information as well as the fluidity of the facial motion influences the activity of the core regions, and Furl et al. (2010) report sensitivity of pSTS and posterior fusiform gyrus to intact versus scrambled facial movement using MEG. Computational modeling and theory suggest that directionality is a key aspect of visual biological motion processing, requiring dedicated neural detectors (Giese and Poggio, 2003). For example, the direction of change from a neutral to a happy face conveys a distinct meaning from the reverse direction, differentially affecting amygdala responses (Sato et al., 2010).

Independent from the emotional direction, the sequence of facial movements during relaxation of an emotional expression is not necessarily the exact reverse of the increase of that expression. Therefore, neural detectors that have been exclusively exposed to natural facial dynamics throughout the lifetime of an observer may respond differentially when exposed to artificially reversed timelines that contain non-canonical temporal sequences.

It is unknown which of the core face processing regions are sensitive to these two independent aspects of directionality, emotional directionality and timeline directionality. From an ecological and physiological point of view, one may expect independent cortical detectors for the two: we are equally frequently exposed to increasing and decreasing emotional expressions, yet they differ in ecological meaning and valence. Facial static emotional content has mostly been shown to increase activity in FFA, pSTS and OFA, along with many other regions of the extended face processing network such as the amygdala, temporal, and prefrontal cortex (Pessoa et al., 2002; Winston et al., 2003) (see for review (Atkinson and Adolphs, 2011)). In contrast, temporal deviations from normal movement trajectories would be expected to affect responses of sequence-specific circuitries thought to be present in the pSTS - either by reducing responses due to suboptimal stimulation, or by enhancing responses as a result of violating predictions (Giese and Poggio, 2003; Rao and Ballard, 1999). We used genuine movie recordings to study these two aspects of facial dynamics in a 2-by-2 factorial design that balanced all visual aspects apart from directionality of motion trajectories. A distracting gender-discrimination task and a rapid event-related design with an unpredictable sequence of stimuli were used to emphasize results related to bottom-up, automatic stimulus processing. Our fMRI results show that dorsal and ventral core face processing regions are sensitive to timeline and emotional directionality.

3.3 Methods

Participants

31 healthy participants with normal or corrected-to-normal vision participated in this study. Data of 27 participants (15 male, mean age 27 ± 4 years, 1 left-handed) entered the final analyses, as a total of 4 participants had to be excluded due to spiking artifacts (2) or exces- sive head-movement (2). The study was conducted according to the declaration of Helsinki and was approved by the local ethics committee of the University of Tübingen. Participants provided written consent prior to participation.

Stimuli

Main experiment

The stimuli of the main experiment included static pictures and short movie clips of faces of eleven actors showing fearful expressions. Movies were recorded prior to the experiment or were obtained from the Video-Face-Database of the MPI Tübingen (Kaulard et al., 2012). All movies were captured in color with the actor placed in front of a black background. Actors showed fearful expressions starting from neutral face, going to peak expression and relaxing back to a neutral expres- sion, and were asked to keep their head still to avoid rigid head movements.

These genuine video recordings were later cut (while maintaining the original frame order) to show either an increase or a decrease of emotional intensity ranging from low to high fear expression or vice versa using VirtualDub (virtualdub.org). The mean durations of the cut movie recordings showing increasing or decreasing fear did statistically not differ (588 ± 139 ms and 680 ± 235 ms respectively). The means of the luminance and of its spatial variance, i.e. root-mean-square (RMS) contrast, for all movies were 96.04 cd/m2 and 109.03 cd/m2, respectively. Duplicates of these movies of increasing and decreasing fear expressions were then reversed in frame order, giving rise to two additional conditions: decreasing and increasing fear in reversed frame order. In total, we obtained four dynamic conditions: increasing and decreasing fear in original frame order, and decreasing and increasing fear in reversed frame order, with 11 exemplars for each. Two static conditions were created using start and end frames of each movie (low and high fear expression, again with 11 exemplars of each). Circular grid-scrambles served as static baseline conditions (Gschwind et al., 2012; Sato et al., 2004). They were obtained by cutting images into tiles of a 10 × 10 grid, and pseudorandomly relocating each tile to a new position that was equidistant to the image center (hence 'circular').

Localizer experiment

For the localizer experiment, neutral and fearful frames of faces from the above videos were contrasted to pictures of houses (kindly provided by Bruno Rossion, http://www.nefy.ucl.ac.be/Face_Categorisation_Lab. htm) as well as to circular grid-scrambles of all pictures. Luminance and RMS contrast of house pictures were adjusted to match those of the faces. Stimuli were back-projected on a screen of 24 × 18 visual degrees, viewed via a tilted mirror and placed centrally, such that stimuli subtended 6 × 9°. All stimuli were presented using Cogent Graphics 1.30 developed by John Romaya at the Wellcome Department of Imaging Neuroscience (http://vislab.ucl.ac.uk/cogent.php) running on MATLAB 2010a on a Windows PC.

Experimental design

Main experiment

The main experiment consisted of seven conditions (Fig. 1B). The four movie conditions displayed an increase or decrease of fear, played either in a forward (natural) or reversed (artificial) frame order as illus- trated in Fig. 1A, yielding a 2 × 2 factorial design (the conditions were: emotion increase in natural frame order, emotion increase in reversed frame order, emotion decrease in natural frame order, and emotion decrease in reversed frame order). The first and last frames were presented for an additional 100 ms to enhance recognizability of the movies by reducing forward- or backward masking effects induced by the isoluminant gray that was shown before and after the stimuli. This resulted in a mean duration of 834 ms for dynamic stimuli. The three static conditions showed start- or end-frames of the movie conditions (low or high fear expression), and grid-scrambled faces, each for 800 ms per trial. Trials of all seven conditions, including static and dynamic conditions, were presented in pseudorandom, history-matched sequences in an event-related design (Fig. 1C), such that every condition preceded equally often all conditions, with seven condition repetitions (49 trials) per run. Inter-stimulus-intervals (ISI) varied randomly between 3000 and 3500 ms (in steps of 125 ms) during which a fixation cross was shown on an isoluminant gray background. The event-related design coupled with the pseudorandomized stimulus sequence eliminated predictability of conditions, therefore avoiding top-down effects of condition-related expectation or attention (at the cost of reduced statistical contrast-efficiency compared to blocked- or pseudo-blocked designs (Liu, 2004)). To ensure vigilance and matched attention across all conditions, participants performed a gender discrimination task, pressing one of two buttons after each trial with the right hand (one subject responded using the left hand). Each run lasted 3.7 min and participants participated in 10 runs. For technical reasons, one subject was scanned with only 8 runs.

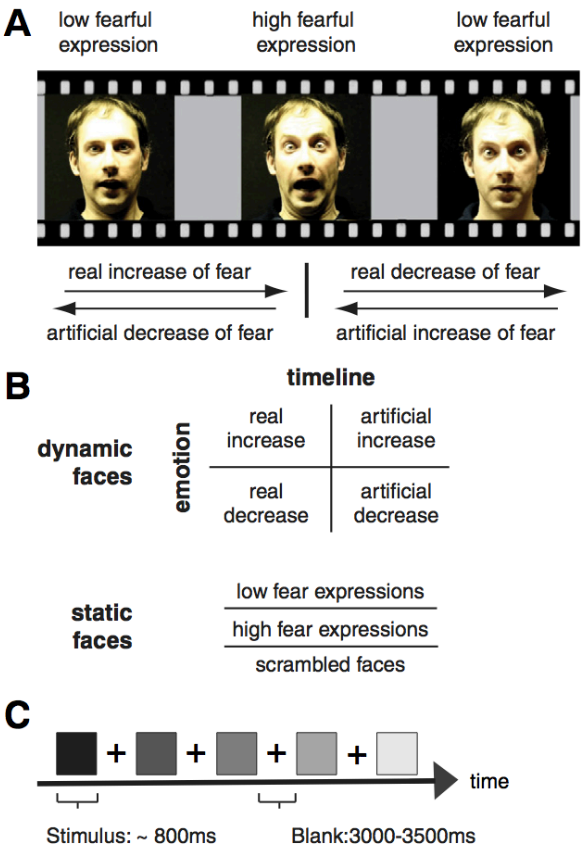


Fig. 1. Stimuli and experimental paradigm.
(A) Illustration of how the four dynamic conditions of increasing and decreasing fearful expressions were obtained in both natural and artificial (reversed) timelines. (B) Overview of the 2-by-2 factorial design of the dynamic conditions with the factors emotion-direction (increasing or decreasing) and timeline (natural or artificial), and of the additional static conditions (movie start- and end-frames, scramble baseline). (C) Timing of the stimulus sequence of the event- related design.

Computerized stimulus ratings

Several visual low-level properties of the movies were quantified using a computer algorithm such that they could account for additional signal variance not related to the high-level properties of interest in the fMRI analysis (Bartels et al., 2008). These properties were the following: the maximal spatial contrast within each movie (RMS normalized by luminance: 1.21 ± 0.30), the temporal contrast (the sum of pixel-wise luminance-changes across the length of the movie: 59.05 ± 39.02 cd/m2), the duration of each movie (see above), and the velocity in the fastest frame pair of each movie (21.95 ± 13.66 pixel/s). These properties were quantified using the methodology described in detail in a prior study (Bartels et al., 2008), and were included as parametric regressors of no interest in the subsequent GLM analysis.

Localizer experiment

Regions of interest (ROIs) related to visual face processing were de- fined using an independent functional block-design localizer lasting 8.9 min. In each block, lasting 16 s, 12 pictures of houses, neutral or fearful faces, or scrambled stimuli were presented. Each picture was shown for 1 s, with ISIs that varied randomly between 240 and 360 ms (in steps of 30 ms) containing a fixation cross on a gray isoluminant background. Blocks of each of the four stimulus categories were repeated 8 times in a history-matched sequence. To ensure vigilance and balanced attention, participants performed a one-back matching task to detect randomly occurring image repetitions that occurred with a frequency of 20%.

Image acquisition

Imaging was performed on a Siemens 3T TIM Trio system (Siemens, Erlangen, Germany) with a 12-channel phased-array head coil. An echoplanar imaging (EPI) sequence was applied to collect T2*-weighted images (EPI) with TR = 2.3 s, TE = 35 ms, flip angle 79°, 33 slices, resulting in a voxel size of 3 × 3 × 3mm. For the main experiment, 98 functional images were acquired in each session, and during the localizer 232 images. In addition, a high-resolution anatomical image was recorded using a T1-weighted MPRAGE sequence yielding 1×1×1mmresolution.

fMRI data preprocessing

Data were processed using SPM5 (www.fil.ion.ucl.ac.uk/spm/). The first four images of each scanning session were discarded as dummy volumes to allow for equilibration of the T1 signal. Data were slice-time corrected for the differences in acquisition time and realigned to the first image to compensate for head motion. The structural image was coregistered to the mean functional image and both structural and functional data were then normalized to a standard brain using the SPM templates. Functional images were smoothed using a Gaussian Kernel with 6 mm full-width at half-maximum (FWHM) for the single- subject analyses and again with 9 mm (resulting in final smoothness of 11 mm) for the group analyses. A high-pass filter of 128 s cut-off was applied to remove low-frequency signal drifts. For two participants only six, or nine, out of the ten recorded sessions were included in the analyses, respectively, due to extensive head movement in the last sessions.

Statistical analysis

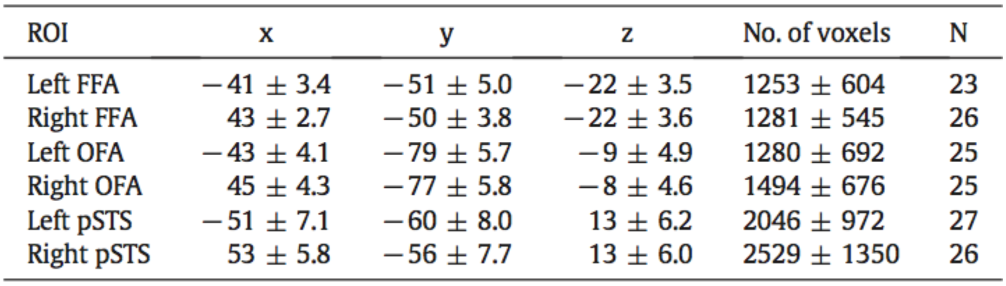
Data of localizer and main experiments of each subject were analyzed in SPM5 using the general linear model (GLM) approach. The design matrix contained one regressor for each condition, modeled using a boxcar convolved with the canonical hemodynamic response function (HRF). Six realignment parameters obtained from the motion correction were included as well a regressor for global signal variance that was orthogonalized against the conditions of interest (Desjardins et al., 2001; Van Dijk et al., 2010). For the main experiment we also included parametric regressors modeling the computationally derived movie- properties described above to account for variance induced by low-level visual features that are independent of the high-level properties of interest in this study.

ROI definition and analysis

Using the localizer experiment, we defined ROIs of the bilateral fusiform face area (FFA) (Kanwisher et al., 1997; McCarthy et al., 1997; Puce et al., 1995; Sergent et al., 1992), the occipital face area (OFA) (Gauthier et al., 2000), and the posterior superior temporal sulcus (pSTS) (Hoffman and Haxby, 2000; Kanwisher et al., 1997; Puce et al., 1998). These three regions are referred to as the “core system” of face processing and are thought to be primarily concerned with visual facial features (Haxby et al., 2000; Ishai et al., 2005; Rossion et al., 2003; Winston et al., 2004), which is why we confined the analysis to them. We used the contrast (faces N houses) to identify each of the three regions, and variable thresholding was applied in the range of $p < 0.001$ (uncorrected) to $p < 0.05$ (FWE corrected) in order to keep the ROIs similar in size across participants (Fox et al., 2009; Jiang et al., 2006; Murray and Wojciulik, 2004) (see Table 1). Where a ROI could not be detected using this method, we attempted to identify it using the contrast "faces versus scramble". This approach is conservative as it has been shown to activate nearly identical coordinates (Fox et al., 2009; Gschwind et al., 2012; Schultz et al., 2012), yet potentially fewer face- specific voxels (Schultz et al., 2012). We also repeated all analyses using ROIs of exactly matched size (using the most significant 50 or 100 voxels of each localizer ROI) (Fox et al., 2009). As this yielded virtually the same results and no systematic differences in significance levels, we report results from the full ROIs as listed in Table 1. For every subject, mean beta estimates were extracted for each ROI and each condition of the main experiment, and the scrambled condition was subtracted from all other conditions prior to further analyses, serving as a common baseline. 2 × 2 ANOVAs with the factors "timeline" (levels: natural, artificial) and "emotion-direction" (levels: increase, decrease) and their interaction were calculated. Results of ROI analyses were corrected for multiple comparison for all ROIs ($n = 6$) using Bonferroni–Holm correction.

Table 1

ROI peak coordinates (in MNI space), number of voxels, and number of ROIs. Provided data are mean ± SD for each of the independently localized core face responsive ROIs.



ROI	x	y	z	No. of voxels	N
Left FFA	−41 ± 3.4	−51 ± 5.0	−22 ± 3.5	1253 ± 604	23
Right FFA	43 ± 2.7	−50 ± 3.8	−22 ± 3.6	1281 ± 545	26
Left OFA	−43 ± 4.1	−79 ± 5.7	−9 ± 4.9	1280 ± 692	25
Right OFA	45 ± 4.3	−77 ± 5.8	−8 ± 4.6	1494 ± 676	25
Left pSTS	−51 ± 7.1	−60 ± 8.0	13 ± 6.2	2046 ± 972	27
Right pSTS	53 ± 5.8	−56 ± 7.7	13 ± 6.0	2529 ± 1350	26

3.4 Results

Responses of core face processing regions were subjected to two- way ANOVAs to determine whether they were differentially sensitive to the factors “emotion-direction” or “timeline” of facial dynamics presented in the 2 × 2 factorial stimulus design. The factor “emotion-direction” tested for differential responses to either increasing or decreasing fear expressions, the factor “timeline” for differential responses to movies played in a forward (natural sequence) or reversed (artificial sequence) frame order. The advantage of the study design was that low- level stimulus properties, static content, and motion properties (except for directionality) were fully counterbalanced across both factors, as half of the stimuli were reversals of the other half (see Fig. 1). This made low-level controls such as time-scrambled movies unnecessary, and allows for a clear attribution of neural signal change to the two factorial stimulus dimensions. Verbal debriefing following the scanning showed as expected that increasing and decreasing fear were easily distinguishable by all participants. Timeline reversal was less obvious: 11 of the 31 participants had noticed that some of the clips were reversals of other clips; 10 were unsure, and 10 did not notice at all. Segregated analyses did not yield differences among ROI responses of these groups of participants (probably also due to the reduced N and the overall small effects). Below results for the whole group are reported.

fMRI results

Regions of the core face processing network, bilateral OFA, FFA and pSTS, were identified using an independent localizer experiment. Table 1 reports their peak coordinates, volumes, and number of participants in which they could be defined. Figs. 2A and 3 show raw responses of each ROI for dynamic and static conditions, respectively.

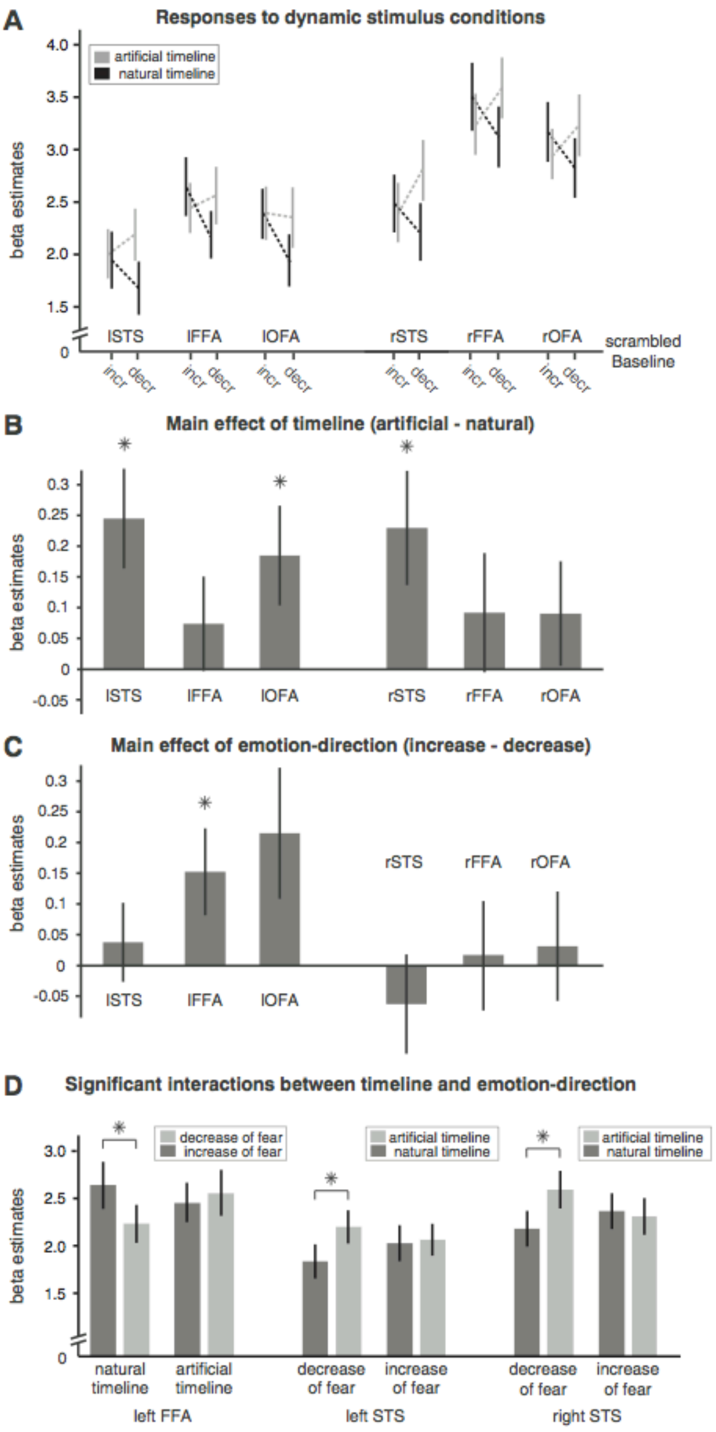


Fig. 2. Main results of ROI responses to dynamic conditions. (A) Raw beta estimates of all core face processing ROIs to the dynamic stimulus conditions. (B) Effects of "timeline", i.e. natural vs. reversed frame order, in all face-responsive ROIs. Presented bars are differences between responses to natural and artificial timelines of left- and right-hemispheric ROIs. (C) Effects of "emotion-direction", i.e. increasing vs. decreasing emotion, in all face-responsive ROIs. Presented bars are differences between responses to increasing and decreasing emotions of left- and right-hemispheric ROIs. (D) Significant interaction effects for left FFA, left and right pSTS. Error bars represent standard errors. *: $p < 0.05$ (2-way ANOVA), Bonferroni–Holm corrected for six ROIs.

First, we consider the factor “timeline”, i.e. responses related to movies played in natural and reversed frame orders, respectively (see Fig. 2B). Among all ROIs, left and right pSTS as well as left OFA had significant main effects for “timeline”, with the following F- and p-values, as well as statistical effect sizes ŋ: left pSTS: F(1,26) = 8.77, p = 0.0065, ŋ2 = 0.251; right pSTS: F(1,25) = 5.90, p = 0.022, ŋ2 = 0.191; and left OFA: F(1,24) = 4.97, p = 0.0354, ŋ2 = 0.172 (all sur- vived Bonferroni-Holm correction for the number of tests performed). pSTS and OFA thus responded stronger to artificial than to natural timelines, whereas FFA had no significant main effect for the factor timeline. Note that the movies entering this contrast were identical apart from their timeline, and that they were also matched in terms of increase or decrease of emotion.

However, the left and right pSTS were not entirely unaffected by the factor “emotion-direction”, in that we also found significant interactions between “timeline” and “emotion-direction” (left pSTS: F(1,26) =4.96, p = 0.0348, ŋ2 = 0.160; right pSTS: F(1,25) = 6.84, p = 0.0149, ŋ2 = 0.215; all surviving Bonferroni–Holm-correction) showing differential responses within timeline only for decreasing fear conditions (see Fig. 2D).

Next, we consider responses related to the factor “emotion-direction” (see Fig. 2C). Only left FFA showed different BOLD responses for “increasing” and “decreasing” fear (F(1,23) = 4.47 p = 0.0455, ŋ2 = 0.162). However, this main effect was driven by a significant interaction between “timeline” and “emotion-direction” (F(1,23) = 7.51, p = 0.0166, ŋ2 = 0.246) in that FFA responses only differed between increasing and decreasing emotion directions within the natural time- line conditions (see Fig. 2D).

FFA's “emotion-direction” effect was therefore entirely driven by the naturally played movies, whereas it did not respond differentially to emotional direction for movies with reversed frame order. This allows us to exclude the alternative explanation, namely that first- or last-frame effects (i.e. low- vs. high static facial expressions) account for the observed response, as this would also have affected responses to the reversed frame-order conditions. First- and last-frame effects can additionally be excluded from accounting for FFA responses, as it did not show any difference between static "low" and "high" fear. In fact, there was no significant response difference between static low and static high fear in any of the ROIs (see Fig. 3).

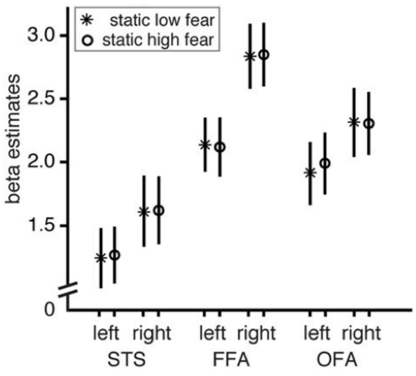


Fig. 3. Raw beta estimates for all core face processing ROIs to static stimulus conditions (i.e. start- and end-frames of the movies). Error bars represent standard errors.

A final analysis further replicated previous work in examining ROI responses to all dynamic versus all static faces (see Fig. 4). While all ROIs responded stronger to dynamic faces ($p < 0.05$, t-tests, Bonferroni corrected), this preference was most pronounced in pSTS, as has been consistently reported in prior studies (Bartels and Zeki, 2004; Fox et al., 2009; Lee et al., 2010; Pitcher et al., 2011; Said et al., 2010; Schultz and Pilz, 2009; Schultz et al., 2012; Trautmann et al., 2009). Note that the advantage of pSTS cannot be accounted for by generally higher responsiveness of pSTS, as its mean response to faces in relation to scramble was similar to that of FFA and OFA (see Fig. 4).

We also performed whole-brain random effect analyses for the key contrasts reported above, but found no activation surviving the usual correction methods, except for the contrast of moving vs. static faces.

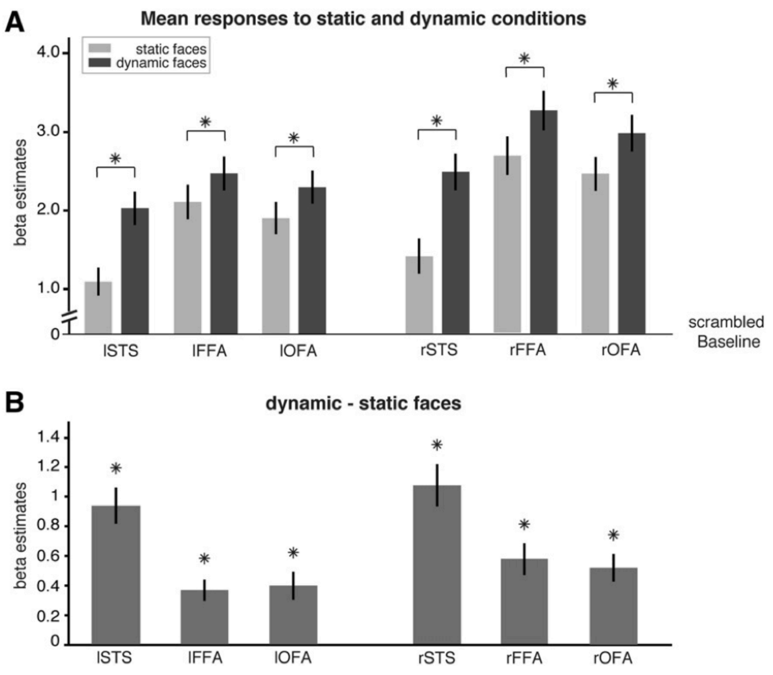


Fig. 4. Mean responses to static and dynamic face stimuli in all face-responsive ROIs. (A): BOLD responses to static and dynamic faces of left- and right-hemispheric ROIs. (B): Differences between responses to dynamic and static faces of left- and right-hemispheric ROIs. Error bars represent standard errors. *: $p < 0.05$ (t-test), Bonferroni corrected.

3.5 Discussion

Directionality is a defining property of all visual movement, and in faces it can have a large impact in transmitting social meaning. We examined whether visual face processing regions differ in their sensitivity to two independent aspects of directionality in facial movement: the direction of emotional expression change, which affects ecological meaning, and the direction of timeline (i.e. natural versus reversed frame order), which exposes sensitivity to prototypical sequences of muscle movements during natural facial expressions (Furl et al., 2010; Giese and Poggio, 2003). Our 2 × 2 factorial design allowed us to determine and attribute brain responses to both factors independently and without confound, as visual low-level properties, static visual content, and motion energy were balanced within and across factors. To our knowledge, this is the first study examining neural effects of reversed movie directions on the processing of dynamic facial expressions. The physical stimulus differences in terms of varied temporal sequences also had perceptual correlates (perceived increasing or decreasing fear; and, perceived by a part of the subjects, the temporal reversal of movies). Thus, our study examined compound neural responses to physical and perceptual effects related to stimuli that differed in temporal sequence but that were matched in static and low-level content.

The present results extend prior knowledge about the functional role of pSTS, OFA and FFA in an important way: FFA showed a sensitivity to the emotion-direction that was timeline-dependent as it only occurred within the natural frame order, and OFA as well as pSTS showed sensitivity to the timeline, with pSTS responses being emotion-direction-dependent as they only occurred for decreased fear. Together, these results reveal interacting temporal sequence sensitive mechanisms that are responsive to both ecological meaning and to prototypical unfolding of facial dynamics.

These findings are of interest to physiologically plausible models of biological movement processing. The differential responses to natural and reversed facial movement trajectories imply either an innate knowledge of typical expression unfolding or a role of experience-dependent plasticity in sequence processing (Giese and Poggio, 2003). The associated sensitivity of core face processing regions to small deviations from natural movement trajectories is relevant in terms of predictive coding theory (Friston, 2005; Rao and Ballard, 1999).

The relatively weak effects observed in this study are probably due to a combination of three factors. First, the stimuli were identical in all respects except for the small differences concerning the timeline only. Second, the task directed attention away from the key feature under study, and third, it was a rapid event-related design. In future studies these aspects could be changed to improve power, potentially at the cost of increasing top-down effects on the observed effects.

OFA and dynamic faces

Haxby et al. (2000) propose a feed-forward model of face processing where

OFA is mainly engaged in early processing steps and provides input to FFA and STS. However, the same authors also emphasize the importance of a 'coordinated participation of multiple regions' (Haxby et al., 2000, p. 231) for different aspects of face processing. In accord with this, evidence from transcranial magnetic stimulation (TMS) suggests a more interactive model that relies on the interplay between the different regions involved (Dzhelyova et al., 2011; Pitcher et al., 2008). For example, OFA is involved not only in early, but also in mid-latency processing of facial properties (Kadosh et al., 2011). Similarly, lesion studies have shown that the presence of FFA and STS alone is not sufficient to discriminate aspects like identity, sex, or emotions of faces (Bouvier and Engel, 2006; Rossion et al., 2003; Steeves et al., 2006). It has therefore been proposed that OFA is one of several interacting nodes that mediate extraction of facial features (Atkinson and Adolphs, 2011), and dynamic faces in particular appear to enhance interactions between OFA and STS (Foley et al., 2012). Our results extend these findings in showing a sensitivity of OFA for deviations from the veridical directionality of dynamic face trajectories.

pSTS and dynamic faces

The pSTS appears to play a particularly important role in analyzing dynamic facial information. However, comparably little is known about what exactly drives the response increase for dynamic faces in pSTS. It is generally sensitive to various forms of biological motion, whether it is of the face (Bartels and Zeki, 2004; Campbell et al., 2001; Fox et al., 2009; LaBar et al., 2003; Puce et al., 1998; Schultz and Pilz, 2009) of point-like walkers (Bonda et al., 1996; Grossman et al., 2000; Peelen et al., 2006), or of human actions and social interactions (Adolphs, 2009; Blake and Shiffrar, 2007; Decety and Grezes, 1999; Schultz et al., 2005). Furthermore, lesions of the pSTS or its deactivation by TMS lead to difficulty in action recognition (Battelli et al., 2003; Grossman et al., 2005; Pavlova et al., 2003; Saygin, 2007; Vaina and Gross, 2004; van Kemenade et al., 2012).

Giese and Poggio (2003) describe a model for processing biological motion where motion pattern neurons integrate information from so-called snapshot-neurons with that from optic-flow-detector neurons. Asymmetric lateral connections between them allow temporal sequence sensitive processing of biological movement. Such distinct circuitries would be expected to react to distinct facial emotional directions. Indeed, a recent study demonstrated a distributed encoding of dynamic facial expressions in pSTS. Voxel patterns encoded seven different dynamic emotional facial expressions, and their similarity structure corresponded to that of perceptual ratings (Said et al., 2010). Similarly, macaque monkey STS voxel pattern also encoded dynamic facial expressions, but generalized poorly to static ones, suggesting that dynamic and static expressions are differentially represented in STS (Furl et al., 2012). Also consistent with the model, electrophysiology in STS has distinguished "snapshot" neurons responsive to static shapes from "motion" neurons responsive to trajectories during viewing of body actions (Vangeneugden et al., 2009).

Action sequences and prediction error

In our experiment, "snapshot" responses were matched, as static content was equal within each factor. That left OFA and pSTS responded more strongly to reversed facial action sequences therefore indicates that additional resources were recruited when trajectories deviated from those typically experienced on a daily basis.

Theories of predictive coding posit exactly that: correct predictions about sensory input lead to a reduction of activity, but to enhanced activity in case of prediction errors (Friston, 2005; Rao and Ballard, 1999). Empirical studies provide evidence compatible with this notion, particularly in context of visual motion (Alink et al., 2010; Bartels et al., 2008; Kanai et al., 2011; Muckli et al., 2005; Murray et al., 2002; Smith and Muckli, 2010; Zaretskaya et al., 2013). The present data provide the first fMRI evidence suggestive of predictive coding in pSTS in context of facial dynamics while keeping static content and non-directional motion properties entirely matched. Prediction error signals are thought to underlie plasticity, and prior fMRI adaptation studies conducted before and after training of biological motion trajectories indeed suggest learning-related plasticity in pSTS (Jastorff et al., 2009).

Two prior studies compared movies of original facial expression changes with their sequence-scrambled counterparts that lacked motion flow and predictability (Furl et al., 2010; Schultz et al., 2012). The observed reduction of pSTS activity with scrambled sequences was attributed to the disruption of motion flow, with only “snapshot” responsive circuitry responding (Schultz et al., 2012). MEG experiments observed a signal modulation in early visual cortex that was interpreted as a prediction-related signal specific to the intact frame-order (Furl et al., 2010).

But what do we know about the temporal unfolding of an emotion? The Facial Action Coding System (FACS) describes prototypical patterns of muscle activations only for peak expressions (Ekman et al., 2002), not entirely considering the temporal unfolding of the emotion. Even if the same Action Units are involved in unfolding and ending an emotion, our data suggest that the order and speed of muscle movement are different, even though these differences remain to be quantified. There is evidence that dynamics in expressions enhance emotion recognition (“dynamic advantage”) (Wehrle et al., 2000). However, artificially altered dynamics have not been systematically studied so far. Detection of consistency or deviation from prototypical trajectories is relevant for social interaction, attribution of intention and credibility, and theory of mind, all functions also associated with the pSTS region (Apperly et al., 2004; Bahnemann et al., 2010; Baron-Cohen et al., 1999; Fletcher et al., 1995; Gallagher et al., 2000; Rilling et al., 2004; Saxe and Kanwisher, 2003; Schultz et al., 2005; Vollm et al., 2006). Interestingly, a region near our left pSTS location has been associated with the processing of complex signals relevant to the communicative significance of other people's behavior (Bahnemann et al., 2010).

Emotion direction specific responses in FFA

FFA was modulated by the “emotion-direction” factor. This is the first evidence showing that FFA can differentiate between increasing and decreasing facial expressions despite similar static content. Note that in contrast to the “timeline” factor (i.e. artificial vs. natural timeline of facial expression), both conditions of the “emotion-direction” factor (i.e. increase or decrease in facial expression) are ecologically equally valid, and are likely to be observed with equal frequency, as each in- crease of facial expression is followed by a decrease. The interpretation of FFA's preference to increasing versus decreasing expression dynamics has to center on factors such as higher social relevance or saliency.

This interpretation gains indirect support from prior studies examining emotional versus neutral expressions. Although FFA has been proposed to be more concerned with invariant aspects of faces, such as identity (Haxby et al., 2000), FFA has previously also been shown to respond more to emotional compared to neutral expressions using static images (Dolan et al., 1996; Gerber et al., 2008; Ishai et al., 2004; Pessoa et al., 2002; Surguladze et al., 2003; Vuilleumier and Pourtois, 2007; Vuilleumier et al., 2001; Winston et al., 2003), and also when conveyed through facial movement (Atkinson et al., 2012). Our result of a main effect in FFA related to emotion direction suggests that FFA's emotion response is driven by behavioral meaning, also when static content is matched. Our study, like most previous ones, cannot distinguish between effects that originate in FFA through feature-detectors, or effects that reflect modulation of FFA related to the conscious percept of stimuli that vary in their emotional valence. However, the previously observed emotion-driven modulation of FFA cannot be fully accounted for by attentional modulation, as it persisted in studies directing attention away from faces or from their expression, as we did using our distractor task in the present study (Pessoa et al., 2002; Vuilleumier and Pourtois, 2007; Vuilleumier et al., 2001).

Our finding that the emotion-direction-related FFA modulation occurred only for natural timeline conditions suggests that either the FFA's response itself, or the input driving it was more sensitive to the correct facial trajectory. This information may be mediated to the FFA by OFA or pSTS that our results show to be sequence sensitive. A potential model may be that the pSTS predominantly forwards sequence-related information from stimuli that match its predictions (i.e. ecologically valid trajectories) and that FFA evaluates their saliency.

Alternatively, the amygdala may provide saliency-related input to FFA (Vuilleumier and Pourtois, 2007). Sato et al. (2010) found the amygdala modulated by dynamic emotion direction. They found no effects in FFA (or pSTS), perhaps because they used linear morph sequences that did not contain natural facial movement trajectories (Sato et al., 2010). Note though that amygdala responses in prior studies tended to be only modulated when linear morph-sequences were used as stimuli, but not with natural expressions (see van der Gaag et al., 2007).

Accounts of the present FFA responses in terms of differential attention to the distinct movie conditions are unlikely, for several reasons. First, it is unclear why certain types of movies used here should attract inherently more attention than others, as all contained matched motion dynamics and matched static content. Also, emotion-specific attention cannot account for the lack of emotion-direction-effect for reversed movies. Second, we used a fast, randomized event-related paradigm that avoided build-up of expectations or mindset related to a given stimulus-category, and our continuous gender-discrimination task was intended to equate attentional vigilance across conditions. Third, and most importantly, generic attention-driven modulation would be expected to modulate all core face processing regions to a similar extent - the observed regional segregation in activity modulation according to the experimental factor cannot be easily accounted for by generic attention effects, unless a similar regional preference to different aspects of facial dynamics is assumed as the one proposed here.

Dynamic and static faces

In contrast to the dynamic conditions, we found no response differ- ences in FFA, OFA or STS to static frames of high and low emotional conditions. This could be due to several reasons. Our paradigm was overall sub-optimal with regards to enhancing signal power related to the stimulus content, as we directed attention to a different feature (gender), and presented stimuli in a statistically inefficient event-related paradigm (Liu, 2004). The fact that we did find significant effects related to our subtle dynamic manipulations that were matched in static content suggests that the dynamic manipulations were comparably more powerful in driving the core face processing regions. A similar conclusion was reached by Trautmann et al. (2009) who similarly found no modulation in core face processing regions between neutral and emotional static faces, but did find emotional modulation for corresponding dynamic facial expressions. Surguladze et al. (2003) report activity increase in the fusiform gyrus for combined 50% and 100% fear expressions compared to neutral, with no involvement of STS or OFA, compatible with our findings. They report pooled activity of both fear intensities, leaving open the possibility that 50% and 100% fear intensity led to similar results. This could be an alternative account for our null finding on static expressions, as we had low fear rather than neutral as comparison to high fear.

Generalization

The present study limited itself to examine responses to fearful expressions, which opens the question to which extent the findings would generalize across other facial expressions. We believe two points are important to consider in this context. First, the “timeline” factor depends entirely on temporal asymmetries in expression dynamics, i.e. the extent to which increase and decrease of an expression follow the same motion trajectory. Therefore, depending on the degree to which other expressions exhibit their temporal asymmetries, their perception and neural responses to the “timeline” factor would be expected to vary. Future studies would be required to examine this. Secondly, the presence of distinct neural responses sensitive to temporal asymmetries on the one hand, and to emotional direction on the other, provides physiological “proof of principle” for the presence of two types of facial motion trajectory sensitive mechanisms that seem to interact at higher processing levels. Both are compatible with predictions from prior modeling work. While it is possible that the anatomical sites vary de- pending on the facial expression, we believe this to be unlikely as the principle of functional specialization would not lead us to expect different sites to be involved depending on the content of the facial expression, just as little as the content of a scene or the identity of a face alter the core face processing networks involved in their processing. Nevertheless, further evidence is required to examine generalization across expressions.

Acknowledgments

We thank Kathrin Kaulard for making movies of facial expressions available. This work was supported by the Centre for Integrative Neuroscience, University of Tübingen, through the German Excellence Initiative (EXC307), and by the Max Planck Society, Germany.

3.6 References

Adolphs, R., 2009. The social brain: neural basis of social knowledge. Annu. Rev. Psychol. 60, 693–716.

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus predictability reduces responses in primary visual cortex. J. Neurosci. Off. J. Soc. Neurosci. 30, 2960–2966.

Ambadar, Z., Schooler, J.W., Cohn, J.F., 2005. Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. Psychol. Sci. 16, 403–410.

Apperly, I.A., Samson, D., Chiavarino, C., Humphreys, G.W., 2004. Frontal and temporo- parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. J. Cogn. Neurosci. 16, 1773–1784.

Atkinson, A.P., Adolphs, R., 2011. The neuropsychology of face perception: beyond simple dissociations and functional selectivity. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 366, 1726–1738.

Atkinson, A.P., Vuong, Q.C., Smithson, H.E., 2012. Modulation of the face- and body- selective visual regions by the motion and emotion of point-light face and body stimuli. NeuroImage 59, 1700–1712.

Bahnemann, M., Dziobek, I., Prehn, K., Wolf, I., Heekeren, H.R., 2010. Sociotopy in the temporoparietal cortex: common versus distinct processes. Soc. Cogn. Affect. Neurosci. 5, 48–58.

Baron-Cohen, S., Ring, H.A., Wheelwright, S., Bullmore, E.T., Brammer, M.J., Simmons, A., Williams, S.C., 1999. Social intelligence in the normal and autistic brain: an fMRI study. Eur. J. Neurosci. 11, 1891–1898.

Bartels, A., Zeki, S., 2004. Functional brain mapping during free viewing of natural scenes. Hum. Brain Mapp. 21, 75–83.

Bartels, A., Zeki, S., Logothetis, N.K., 2008. Natural vision reveals regional specialization to local motion and to contrast-invariant, global flow in the human brain. Cereb. Cortex 18, 705–717.

Battelli, L., Cavanagh, P., Thornton, I.M., 2003. Perception of biological motion in parietal patients. Neuropsychologia 41, 1808–1816.

Blake, R., Shiffrar, M., 2007. Perception of human motion. Annu. Rev. Psychol. 58, 47–73. Bonda, E., Petrides, M., Ostry, D., Evans, A., 1996. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. J. Neurosci. 16, 3737–3744.

Bouvier, S.E., Engel, S.A., 2006. Behavioral deficits and cortical damage loci in cerebral achromatopsia. Cereb. Cortex 16, 183–191.

Campbell, R., 1992. The neuropsychology of lipreading. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 335, 39–45. Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J.,

Brammer, M.J., David, A.S., 2001. Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). Brain Res. Cogn. Brain Res. 12, 233–243.

Decety, J., Grezes, J., 1999. Neural mechanisms subserving the perception of human actions. Trends Cogn. Sci. 3, 172–178.

Desjardins, A.E., Kiehl, K.A., Liddle, P.F., 2001. Removal of confounding effects of global signal in functional MRI analyses. NeuroImage 13, 751–758.

Dolan, R.J., Fletcher, P., Morris, J., Kapur, N., Deakin, J.F., Frith, C.D., 1996. Neural activation during covert processing of positive emotional facial expressions. NeuroImage 4, 194–200.

Dzhelyova, M.P., Ellison, A., Atkinson, A.P., 2011. Event-related repetitive TMS reveals distinct, critical roles for right OFA and bilateral posterior STS in judging the sex and trustworthiness of faces. J. Cogn. Neurosci. 23, 2782–2796.

Ekman, P., Friesen, W.V., Hager, J.C., 2002. The Facial Action Coding System, 2nd ed. Research Nexus eBook, Salt Lace City, UT.

Fletcher, P.C., Happé, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S.J., Frith, C.D., 1995. Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition 57, 109–128.

Foley, E., Rippon, G., Thai, N.J., Longe, O., Senior, C., 2012. Dynamic facial expressions evoke distinct activation in the face perception network: a connectivity analysis study. J. Cogn. Neurosci. 24, 507–520. Fox, C.J.,

Iaria, G., Barton, J.J., 2009. Defining the face processing network: optimization of the functional localizer in fMRI. Hum. Brain Mapp. 30, 1637–1651.

Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 360, 815–836.

Furl, N., van Rijsbergen, N.J., Kiebel, S.J., Friston, K.J., Treves, A., Dolan, R.J., 2010. Modulation of perception and brain activity by predictable trajectories of facial expressions. Cereb. Cortex 20, 694–703.

Furl, N., Garrido, L., Dolan, R.J., Driver, J., Duchaine, B., 2011. Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. J. Cogn. Neurosci. 23, 1723–1740.

Furl, N., Hadj-Bouziane, F., Liu, N., Averbeck, B.B., Ungerleider, L.G., 2012. Dynamic and static facial expressions decoded from motion-sensitive areas in the macaque monkey. J. Neurosci. Off. J. Soc. Neurosci. 32, 15952–15962.

Gallagher, H.L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., Frith, C.D., 2000. Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and non- verbal tasks. Neuropsychologia 38, 11–21.

Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000. The fusiform "face area"" is part of a network that processes faces at the individual level. J. Cogn. Neurosci. 12, 495–504.

Gerber, A.J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J., Peterson, B.S., 2008. An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. Neuropsychologia 46, 2129–2139.

Giese, M.A., Poggio, T., 2003. Neural mechanisms for the recognition of biological movements. Nat. Rev. Neurosci. 4, 179–192.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., Blake, R., 2000. Brain areas involved in perception of biological motion. J. Cogn. Neurosci. 12, 711–720.

Grossman, E.D., Battelli, L., Pascual-Leone, A., 2005. Repetitive TMS over posterior STS disrupts perception of biological motion. Vis. Res. 45, 2847–2853.

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., Vuilleumier, P., 2012. White- matter connectivity between face-responsive regions in the human brain. Cereb. Cortex 22, 1564–1576.

Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. Trends Cogn. Sci. 4, 223–233.

Hill, H., Johnston, A., 2001. Categorizing sex and identity from the biological motion of faces. Curr. Biol. 11, 880–885.

Hoffman, E.A., Haxby, J.V., 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nat. Neurosci. 3, 80–84.

Ishai, A., Pessoa, L., Bikle, P.C., Ungerleider, L.G., 2004. Repetition suppression of faces is modulated by emotion. Proc. Natl. Acad. Sci. U. S. A. 101, 9827–9832.

Ishai, A., Schmidt, C.F., Boesiger, P., 2005. Face perception is mediated by a distributed cortical network. Brain Res. Bull. 67, 87–93.

Jastorff, J., Kourtzi, Z., Giese, M.A., 2009. Visual learning shapes the processing of complex movement stimuli in the human brain. J. Neurosci. Off. J. Soc. Neurosci. 29, 14026–14038.

Jiang, X., Rosen, E., Zeffiro, T., Vanmeter, J., Blanz, V., Riesenhuber, M., 2006. Evaluation of a shape-based model of human face discrimination using FMRI and behavioral techniques. Neuron 50, 159–172.

Kadosh, K.C., Walsh, V., Kadosh, R.C., 2011. Investigating face-property specific processing in the right OFA. Soc. Cogn. Affect. Neurosci. 6, 58–65.

Kanai, R., Carmel, D., Bahrami, B., Rees, G., 2011. Structural and functional fractionation of right superior parietal cortex in bistable perception. Curr. Biol. 21, R106–R107.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Kaul, C., Rees, G., Ishai, A., 2011. The gender of face stimuli is represented in multiple regions in the human brain. Front. Hum. Neurosci. 4, 238.

Kaulard, K., Cunningham, D.W., Bulthoff, H.H., Wallraven, C., 2012. The MPI facial expression database — a validated database of emotional and conversational facial expressions. PLoS One 7, e32321.

Kilts, C.D., Egan, G., Gideon, D.A., Ely, T.D., Hoffman, J.M., 2003. Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. NeuroImage 18, 156–168.

Kriegeskorte, N., Formisano, E., Sorger, B., Goebel, R., 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc. Natl. Acad. Sci. U. S. A. 104, 20600–20605.

LaBar, K.S., Crupain, M.J., Voyvodic, J.T., McCarthy, G., 2003. Dynamic perception of facial affect and identity in the human brain. Cereb. Cortex 13, 1023–1033.

Lee, L.C., Andrews, T.J., Johnson, S.J., Woods, W., Gouws, A., Green, G.G., Young, A.W., 2010. Neural responses to rigidly moving faces displaying shifts in social attention investigated with fMRI and MEG. Neuropsychologia 48, 477–490.

Liu, T.T., 2004. Efficiency, power, and entropy in event-related fMRI with multiple trial types. Part II: design of experiments. NeuroImage 21, 401–413.

McCarthy, G., Puce, A., Gore, J.C., Allison, T., 1997. Face-specific processing in the human fusiforrn gyms. J. Cogn. Neurosci. 9, 605–610.

Muckli, L., Kohler, A., Kriegeskorte, N., Singer, W., 2005. Primary visual cortex activity along the apparent-motion trace reflects illusory perception. PLoS Biol. 3, e265.

Murray, S.O., Wojciulik, E., 2004. Attention increases neural selectivity in the human lateral occipital complex. Nat. Neurosci. 7, 70–74.

Murray, S.O., Kersten, D., Olshausen, B.A., Schrater, P., Woods, D.L., 2002. Shape perception reduces activity in human primary visual cortex. Proc. Natl. Acad. Sci. U. S. A. 99, 15164–15169.

Natu, V., Raboy, D., O'Toole, A.J., 2011. Neural correlates of own- and other-race face perception: spatial and temporal response differences. NeuroImage 54, 2547–2555.

Nestor, A., Plaut, D.C., Behrmann, M., 2011. Unraveling the distributed neural code of facial iden- tity through spatiotemporal pattern analysis. Proc. Natl. Acad. Sci. U. S. A. 108, 9998–10003.

Nummenmaa, L., Calder, A.J., 2009. Neural mechanisms of social attention. Trends Cogn. Sci. 13, 135–143. O'Toole, A.J., Roark, D.A., Abdi, H., 2002. Recognizing moving faces: a psychological and neural synthesis. Trends Cogn. Sci. 6, 261–266.

Pavlova, M., Staudt, M., Sokolov, A., Birbaumer, N., Krageloh-Mann, I., 2003. Perception and production of biological movement in patients with early periventricular brain lesions. Brain 126, 692–701.

Peelen, M.V., Wiggett, A.J., Downing, P.E., 2006. Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. Neuron 49, 815–822.

Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G., 2002. Neural processing of emotional faces requires attention. Proc. Natl. Acad. Sci. U. S. A. 99, 11458–11463.

Pitcher, D., Garrido, L., Walsh, V., Duchaine, B.C., 2008. Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. J. Neurosci. Off. J. Soc.

Pitcher, D., Dilks, D.D., Saxe, R.R., Triantafyllou, C., Kanwisher, N., 2011. Differential selectivity for dynamic versus static information in face-selective cortical regions. NeuroImage 56, 2356–2363.

Puce, A., Perrett, D., 2003. Electrophysiology and brain imaging of biological motion. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 358, 435–445.

Puce, A., Allison, T., Gore, J.C., McCarthy, G., 1995. Face-sensitive regions in human extrastriate cortex studied by functional MRI. J. Neurophysiol. 74, 1192–1199.

Puce, A., Allison, T., Bentin, S., Gore, J.C., McCarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. J. Neurosci. Off. J. Soc. Neurosci. 18, 2188–2199.

Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpre- tation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87.

Rilling, J.K., Sanfey, A.G., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2004. The neural correlates of theory of mind within interpersonal interactions. NeuroImage 22, 1694–1703.

Rossion, B., Caldara, R., Seghier, M., Schuller, A.M., Lazeyras, F., Mayer, E., 2003. A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. Brain 126, 2381–2395.

Said, C.P., Moore, C.D., Engell, A.D., Todorov, A., Haxby, J.V., 2010. Distributed representa- tions of dynamic facial expressions in the superior temporal sulcus. J. Vis. 10 (11-11).

Said, C.P., Haxby, J.V., Todorov, A., 2011. Brain systems for assessing the affective value of faces. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 366, 1660–1670.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., Matsumura, M., 2004. Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. Brain Res. Cogn. Brain Res. 20, 81–91.

Sato, W., Kochiyama, T., Yoshikawa, S., 2010. Amygdala activity in response to forward versus backward dynamic facial expressions. Brain Res. 1315, 92–99.

Saxe, R., Kanwisher, N., 2003. People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. NeuroImage 19, 1835–1842.

Saygin, A.P., 2007. Superior temporal and premotor brain areas necessary for biological motion perception. Brain 130, 2452–2461.

Schultz, J., Pilz, K.S., 2009. Natural facial motion enhances cortical responses to faces. Exp. Brain Res. 194, 465–475.

Schultz, J., Friston, K.J., O'Doherty, J., Wolpert, D.M., Frith, C.D., 2005. Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. Neuron 45, 625–635.

Schultz, J., Brockhaus, M., Bulthoff, H.H., Pilz, K.S., 2012. What the human brain likes about facial motion. Cereb. Cortex 12, 1167–1178.

Sergent, J., Ohta, S., Macdonald, B., 1992. Functional anatomy of face and object processing. Brain 115, 15–36.

Smith, F.W., Muckli, L., 2010. Nonstimulated early visual areas carry information about surrounding context. Proc. Natl. Acad. Sci. U. S. A. 107, 20099–20103.

Steeves, J.K., Culham, J.C., Duchaine, B.C., Pratesi, C.C., Valyear, K.F., Schindler, I., Humphrey, G.K., Milner, A.D., Goodale, M.A., 2006. The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. Neuropsychologia 44, 594–609.

Steeves, J., Dricot, L., Goltz, H.C., Sorger, B., Peters, J., Milner, A.D., Goodale, M.A., Goebel, R., Rossion, B., 2009. Abnormal face identity coding in the middle fusiform gyrus of two brain-damaged prosopagnosic patients. Neuropsychologia 47, 2584–2592.

Surguladze, S.A., Brammer, M.J., Young, A.W., Andrew, C., Travis, M.J., Williams, S.C., Phillips, M.L., 2003. A preferential increase in the extrastriate response to signals of danger. NeuroImage 19, 1317–1328.

Thornton, I.M., Kourtzi, Z., 2002. A matching advantage for dynamic human faces. Perception 31, 113–132.

Trautmann, S.A., Fehr, T., Herrmann, M., 2009. Emotions in motion: dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. Brain Res. 1284, 100–115.

Vaina, L.M., Gross, C.G., 2004. Perceptual deficits in patients with impaired recognition of biological motion after temporal lobe lesions. Proc. Natl. Acad. Sci. U. S. A. 101, 16947–16951.

van der Gaag, C., Minderaa, R.B., Keysers, C., 2007. The BOLD signal in the amygdala does not differentiate between dynamic facial expressions. Soc. Cogn. Affect. Neurosci. 2, 93–103.

Van Dijk, K.R., Hedden, T., Venkataraman, A., Evans, K.C., Lazar, S.W., Buckner, R.L., 2010. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. J. Neurophysiol. 103, 297–321.

van Kemenade, B.M., Muggleton, N., Walsh, V., Saygin, A.P., 2012. Effects of TMS over premotor and superior temporal cortices on biological motion perception. J. Cogn. Neurosci. 24, 896–904.

Vangeneugden, J., Pollick, F., Vogels, R., 2009. Functional differentiation of macaque visual temporal cortical neurons using a parametric action space. Cereb. Cortex 19, 593–611.

Vollm, B.A., Taylor, A.N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J.F., Elliott, R., 2006. Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. NeuroImage 29, 90–98.

Vuilleumier, P., Pourtois, G., 2007. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. Neuropsychologia 45, 174–194.

Vuilleumier, P., Armony, J.L., Driver, J., Dolan, R.J., 2001. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30, 829–841.

Wehrle, T., Kaiser, S., Schmidt, S., Scherer, K.R., 2000. Studying the dynamics of emotional ex- pression using synthesized facial muscle movements. J. Pers. Soc. Psychol. 78, 105–119.

Winston, J.S., O'Doherty, J., Dolan, R.J., 2003. Common and distinct neural responses during direct and incidental processing of multiple facial emotions. NeuroImage 20, 84–97.

Winston, J.S., Henson, R.N., Fine-Goulden, M.R., Dolan, R.J., 2004. fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. J. Neurophysiol. 92, 1830–1839.

Zaretskaya, N., Anstis, S., Bartels, A., 2013. Parietal cortex mediates conscious perception Neurosci. 28, 8929–8933. of illusory gestalt. J. Neurosci. Off. J. Soc. Neurosci. 33, 523–531.

4 Perception of temporal asymmetries in dynamic facial expressions

This chapter has been reproduced from an article published in Frontiers in Psychology: Reinl, M., & Bartels, A. (2015) Perception of temporal asymmetries in dynamic facial expressions. *Frontiers in Psychology, 6*, 1107.

4.1 Abstract

In the current study we examined whether timeline-reversals and emotional direction of dynamic facial expressions affect subjective experience of human observers. We recorded natural movies of faces that increased or decreased their expressions of fear, and played them either in the natural frame order or reversed from last to first frame (reversed timeline). This led to four conditions of increasing or decreasing fear, either following the natural or reversed temporal trajectory of facial dynamics. This 2-by-2 factorial design controlled for visual low-level properties, static visual content, and motion energy across the different factors. It allowed us to examine perceptual consequences that would occur if the timeline trajectory of facial muscle movements during the increase of an emotion are not the exact mirror of the timeline during the decrease. It additionally allowed us to study perceptual differences between increasing and decreasing emotional expressions. Perception of these time-dependent asymmetries have not yet been quantified. We found that three emotional measures, emotional intensity, artificialness of facial movement, and convincingness or plausibility of emotion portrayal, were affected by timeline-reversals as well as by the emotional direction of the facial expressions. Our results imply that natural dynamic facial expressions contain temporal asymmetries, and show that deviations from the natural timeline lead to a reduction of perceived emotional intensity and convincingness, and to an increase of perceived artificialness of the dynamic facial expression. In addition, they show that decreasing facial expressions are judged as less plausible than increasing facial expressions. Our findings are of relevance for

both, behavioral as well as neuroimaging studies, as processing and perception are influenced by temporal asymmetries.

4.2 Introduction

Facial expressions are dynamic by nature. It is therefore not surprising that facial motion is a fundamental source of information for social interactions. The importance of motion for face perception has been recognized many years ago (Bassili, 1978; Tomkins, 1982), and several different lines of research have demonstrated that facial motion has indeed facilitative effects on a variety of perceptual and psychological processes. Humphreys et al. (1993) described a patient with visual object agnosia who failed to recognize identity and emotions of static faces, but performed at normal levels when viewing dynamic faces. More recently, facial dynamics have been shown to increase performance on emotion ratings also in healthy participants (Harwood et al., 1999; Wehrle et al., 2000; Ambadar et al., 2005; Biele and Grabowska, 2006; Weyers et al., 2006; Cunningham and Wallraven, 2009) as well as the encoding and recognition of facial identity (Hill and Johnston, 2001; O'Toole et al., 2002; Thornton and Kourtzi, 2002; Knappmeyer et al., 2003; Pilz et al., 2005; Lander et al., 2006). Lederman et al. (2007) showed that dynamic stimuli can improve haptic recognition of emotional faces, and facial dynamics have also been shown to affect physiological automatic responses in that they enhance facial mimicry and affected physiological measures of arousal rates such as heart rate or skin conductance (Simons et al., 1999; Weyers et al., 2006; Sato and Yoshikawa, 2007a).

However, only little is known about possible reasons that drive the perceptual advantage of dynamic face stimuli. Dynamic displays can be described as a series of static images that provide an increased amount of information to the observer, which could be the underlying cause for the dynamic advantage. In Ambadar et al. (2005) tested this hypothesis by presenting either static, multi-static (frames of dynamic stimuli separated by masks to disrupt the percept of coherent motion) or dynamic stimuli of emotional expressions. If an increase of static information would account for the dynamic superiority, one would expect both the multi-static and the dynamic condition to lead to better recognition results. However, this was not the case. Only the dynamic condition improved recognition rates. This suggests that dynamic sequences carry a distinct source of information that is not present in additional static cues. In line with this, Lander and Bruce (2004) observed impairments in identity recognition for scrambled, reversed and decelerated dynamic face movies. Pollick et al. (2003) found that spatial exaggeration of motion trajectories had a substantial effect on recognition rates and intensity ratings of different emotions while temporal variations only lead to small effects on emotion perception. Despite this, humans were shown to be highly sensitive in detecting small changes in the time course of facial movement trajectories (Dobs et al., 2014) and able to reproduce the temporal order of facial expressions from a scrambled set of photographs (Edwards, 1998).

Imaging studies have shown that brain regions responsive to static faces increased their activity in response to facial motion, even when attention was distracted, suggesting that additional neural processes are recruited for processing dynamic compared to static faces (Kilts et al., 2003; Labar et al., 2003; Sato et al., 2004; Fox et al., 2009; Schultz and Pilz, 2009; Trautmann et al., 2009). fMRI also showed a sensitivity of face processing brain regions to the fluidity of facial motion (Schultz et al., 2013), and MEG revealed changes in neural activation for scrambled versus correct-order facial expressions using (Furl et al., 2010). Finally, patient studies showed a dissociation between impairments of static and dynamic facial expressions: PS, a patient with acquired prosopagnosia was impaired in categorizing static facial expressions, but performed normal in categorizing dynamic facial expressions (Richoz et al., 2015).

In sum, it appears that the dynamics, the timing, and the correct temporal sequence of dynamic facial expression changes are crucial for the dynamic face advantage. Thus, directionality is a key aspect in dynamic face processing and perception. This has also been corroborated by computational modeling and theory of visual biological motion processing (Giese and Poggio, 2003). For faces, the importance of directionality is easily illustrated in the example where the direction of change from a neutral to an emotional facial expression (i.e., increasing fear) carries a different ecological meaning than the reversed direction (i.e., relaxing from fear). Increasing fear could, e.g., signal approaching danger to an external observer, whereas relaxing from fear the opposite, even though the average static face information is identical in both conditions. Correspondingly, prior experiments found that the direction of emotional change had perceptual effects, referred to as representational momentum (Freyd and Finke, 1984; Finke and Freyd, 1985). These studies showed differences in subjective rating of facial emotion, e.g., when the intensity of a neutral facial emotion was rated when it was the end-frame of a movie clip starting with a happy or a sad facial expression (Yoshikawa and Sato, 2008; Jellema et al., 2011; Marian and Shimamura, 2013).

Another temporal instance of directionality has received less attention in the past: the sequence of facial movements during relaxation of an emotional expression may not be the exact reverse of the increase of that expression. In prior studies, activation time courses of facial action units showed temporal asymmetries during basic emotional expressions such as happiness or fear (Dobs et al., 2014; Jack et al., 2014). Also, the information content graspable from the face evolves over time: while the earliest components of facial expressions allow for a crude differentiation of approach versus avoidance, the later components signal socially more complex categories (Jack et al., 2014). We hypothesize that unless the temporal evolvement of increasing and relaxing facial expressions is exactly the same, reversing the direction will be perceived differently by a human observer. Such differences will reflect the presence of temporal asymmetry between the two timecourses. In a previous fMRI study we found that face selective regions responded differentially to natural and reversed timelines of dynamic facial expressions, even if controlled for expression direction (i.e., increase and decrease). Since the corresponding static start- and end-frames of these movies did not elicit distinct responses, these brain regions must have been differentially activated due to asymmetries in facial dynamics (Reinl and Bartels, 2014). In addition, we also found neural effects of the expression direction (increasing versus decreasing), partly independent from the timeline manipulation. In the present study we aimed to examine corresponding behavioral effects, i.e., subjective perceptual consequences of natural versus reversed trajectories. To our knowledge, it is not known whether behavioral judgments of human observers are sensitive to these asymmetries and if, how it influences the evaluation of the facial emotions.

Fear is one of the prototypical expressions of high ecological importance that needs to be transmitted and recognized rapidly in order to act efficiently as a warning sign to peers. We therefore chose this a first expression to examine temporal asymmetry. Even though temporal asymmetry should also be studied in a range of other expressions, we would expect similar effects given that they play a role in one of the prototypical expressions.

We first quantified physical motion in our natural face stimuli in order to test whether temporal asymmetries do exist in our stimulus set. Stimuli were genuine movie recordings of increasing and decreasing fearful expressions. These were presented in the natural forward frame order as well as reversed. This led to a 2-by- 2 factorial design (increasing versus decreasing facial expression, and natural versus reversed frame order). This allowed us to study behavioral effects of timeline reversal as well as emotional directionality. We hypothesized that the visual system is used to certain natural temporal asymmetries. Deviations, such as induced by timeline reversal, would be reflected in a decrease of perceived emotional intensity, as reversed timelines lead to an atypical unfolding of the emotion. Second, we hypothesized that reversed timelines lead to an increase of perceived artificialness, and to a decrease of convincingness or plausibility of the emotion portrayal.

4.3 Materials and Methods

Participants

Data were obtained in two separate sessions. The first session was conducted with 28 caucasian participants (15 male, mean age 27 ± 4 years, 1 left-handed). A second session was conducted a few months later to extend the results from the first session. Unfortunately, it was not possible to re-test all subjects of the first session, so the follow-up session included only 19 of the previous 28 subjects. Subjects were healthy with normal or corrected- to-normal vision. The study was conducted according to the declaration of Helsinki and was approved by the local ethics committee of the University of Tübingen. Participants provided written consent prior to participation.

Stimuli and Procedure

Stimuli consisted of colored short movie clips of eleven caucasian actors portraying fearful facial expressions. Seven movies were recorded prior to the experiment, four movies were selected from the Video-Face-Database of the MPI Tübingen (Kaulard et al., 2012). Actors were asked to show fearful expressions, starting from a neutral face, going to peak expression and relaxing back to a neutral expression. They were asked to keep their head still to minimize rigid head movements. To improve validity of the expression, actors were told to imagine a fearful situation while posing the expression. From every actor several repetitions were recorded. Movies were then selected by visual inspection. Criteria were a recognizable fearful expression with a clear increase and decrease as well no or only little head movement and no excessive eye blinking. Recordings were cut to show either an increase or a decrease of emotional intensity ranging from low to high fear expression or vice versa using VirtualDub (virtualdub.org). Subsequent to this, head-motion was removed by calculating the point-of-gravity (based on luminance values) for each frame and re-centering each frame to its mid-point. The movies were cut at the apex of the expression. The resulting mean durations of both emotional directions (mean and sd: 581 ± 145 ms and 643 ± 245 ms respectively) did not differ statistically [$t(10) = -0.90$, $p = 0.39$]. The means of luminance and of spatial variance for all movies were 96.04 cd/m2 luminance and 109.03 cd/m2 root-mean-square (RMS) contrast, respectively.

Movies were presented in original and in reversed frame order, giving rise to four conditions: increasing and decreasing fear in original frame order (natural timeline), and decreasing and increasing fear in reversed frame order (reversed timeline), with 11 exemplars for each condition (see Figure 1). Subjects were placed in front of a computer monitor and every movie of each condition was presented to them once in a random sequence that was counterbalanced across subjects. Each presentation was followed by a visual presentation of a scale for rating purposes as described below. Subjects had no time limit for their responses. After their response there was a fixation cross of one second duration before the next stimulus appeared on screen.

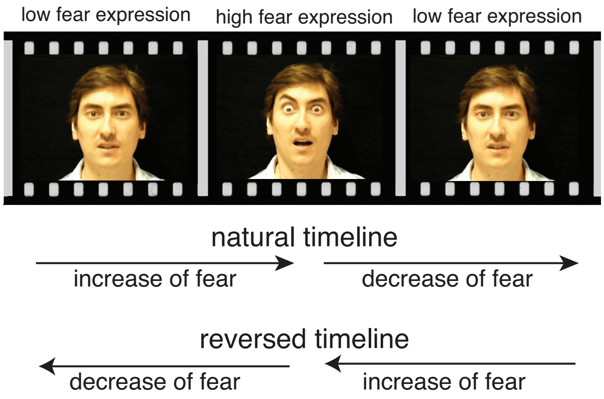


FIGURE 1 | Illustration of stimulus material. The four conditions of the experiment were created by playing increasing and decreasing fearful face movies both in a natural as well as in a reversed frame order.

Quantification of Physical Motion in Facial Trajectories

In order to test - in a very basic fashion - whether the dynamic facial expressions in our dataset did indeed contain temporal asymmetries in their motion content, we subjected our stimuli to a computational optic flow algorithm that has been shown to be physiologically plausible (Bülthoff et al., 1989) and that we have successfully used in the past to relate motion in natural movies to brain activation in visual motion regions (Bartels et al., 2008). This algorithm tracks local motion vectors across each frame-pair. Local motion was estimated in a 20-by-20 grid (resulting in 400 vectors) evenly spaced across each frame. The sum of the lengths of all local motion vectors for each frame-pair yielded a time-series of motion intensity for each of the movies. These motion timelines were interpolated to an equal length of 40 time-points (corresponding to 60 Hz for 600 ms long movies), and range- normalized such that all values fell between 0 and 1 for each movie. Two tests were applied to quantify temporal asymmetry. First, we tested for temporal asymmetry within the same movie, i.e., we subtracted motion intensity timelines of forward from reversed movie, and averaged the absolute differences across the different movies, for every time-point. This yielded a time-series of motion difference. This was done separately for increasing fear expression movies (natural vs. reversed timeline) and for decreasing fear movies (natural vs. reversed timelines). Second, we performed the same procedure across increasing fear movies (original timeline) versus reversed timeline for decreasing fear movies, again performed pair-wise within the same actor and summed across actors. T-tests were then applied for every timepoint and Bonferroni-corrected for the number of timepoints (n = 40) to identify timepoints where timelines were significantly different between forward and reversed timelines, within or across movies, respectively.

Rating of Emotional Intensity

In the first session, subjects were asked to rate the amount of fear presented in the movies on a scale from 1 (low fear) to 6 (high fear). Subjects were instructed to indicate the maximum intensity or strength of fear displayed by the facial expression, regardless of the quality or naturalness of the acting.

Rating of Artificialness and Convincingness

After the intensity rating of the first session revealed significant effects regarding timeline changes, we decided to extend the investigation of perceptual effects and performed a second rating session. In particular, we reckoned apart from intensity, also convincingness and perceived artificialness could be affected by the manipulation. Thus, subjects were subsequently asked to rate two more measures: the artificialness of the actor performance as well as how convincing fear was portrayed by the actors. The reason why we added these two additional measures to our behavioral tests was the following. We assumed that playing movies backward would make the facial dynamics appear unusual, which is tested by the rating of artificialness. Subjects were instructed to rate to which extent they perceived the movement of the face as strange in any way, independent from the quality of the emotion portrayal, its genuineness (i.e., was the actor really feeling the emotion he or she is portraying), or its intensity. After each presentation of a stimuli, they were asked to answer the question "As how natural would you describe the movie?" on scale from 1 ("very natural") to 8 ("very artificial").

Second, we hypothesized that the degree to which a dynamic facial expression is judged as plausible emotion portrayal may be affected by our manipulations as well. Note that this rating, which we refer to as "convincingness," does not necessarily need to be coupled to the artificialness rating, nor to the intensity of the displayed emotion. For example, a strong emotional expression can be perceived as entirely acted and non-genuine (i.e., the actor was not really feeling the emotion). Equally, a well-acted emotional expression can be non-genuine but highly plausible/convincing. The distinction between these ratings has also been discussed previously (Sato and Yoshikawa, 2004; Krumhuber and Kappas, 2005). We asked our subjects to evaluate how convincing or plausible the emotion was portrayed by the actors regardless of whether it seemed to be genuine or not. After each presentation of a stimuli, they were asked to answer the question "How convincing was the emotion fear portrayed?" on a scale from 1 ("very convincing") to 8 ("not convincing"). Note, that for a better understanding the scores have been reversed in the result figure (Figure 3, results are presented from 1 "not convincing" to 8 "very convincing").

4.4 Results

First, a quantification of motion intensity over time was carried out for each movie in order to obtain a very basic measure of asymmetry of motion content over time. Using the obtained motion-intensity timelines of each movie three tests were carried out. First, we tested for temporal asymmetry within natural recordings of increasing fear, and separately for those of decreasing fear. Figure 2 shows the average timeline of motion intensity for increasing and decreasing fear, respectively.

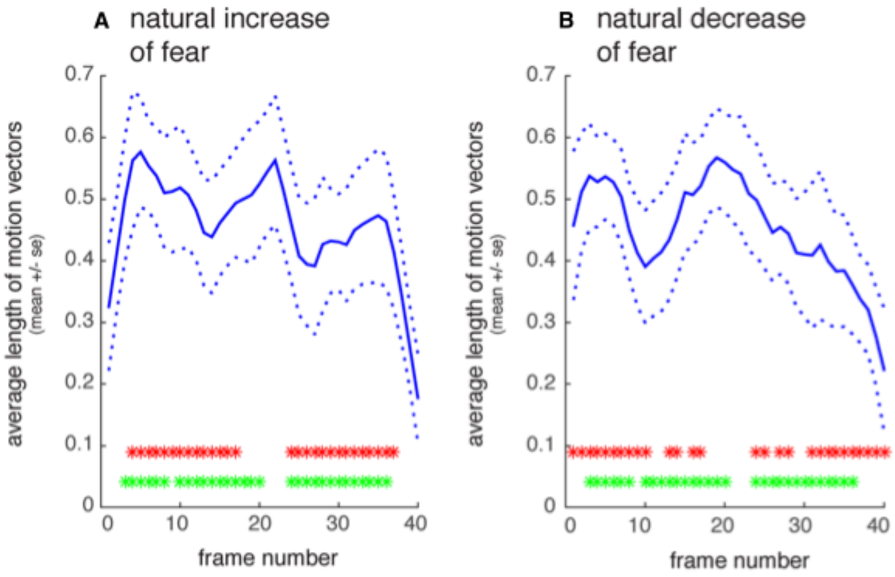


FIGURE 2 | Quantification of motion intensity over time. (A) Natural recordings of increasing fear, averaged across 11 actors. (B) Natural recordings of decreasing fear. The solid line shows the normalized average length of local motion vectors that were estimated for each frame-pair; dotted line: SE The upper row of asterisks shows time-points where the natural and reversed timelines differ significantly ($p < 0.05$, Bonferroni-corrected for 40 frames) in motion intensity. The lower row of asterisks shows timepoints where increasing fear movies differ in motion from reversed decreasing fear movies (same correction as above). X-axis denotes time in movie-frames (1/60s), y-axis normalized motion intensity.

For each of these two emotion directions, we subtracted each individual movie motion timeline from its reversed counterpart to test for temporal asymmetry within a given movie category. For the most part the natural and reversed timelines differed, i.e., each movie category turned out to have asymmetric motion timelines (increasing fear: $T(11) = 8.05$, $p < 0.001$; decreasing fear:

T(11) = 10.06, p < 0.001). Second, we tested whether increasing fear movies were matched in their motion timeline by reversed timelines of decreasing fear movies. Again, for the most part the timelines differed, indicating asymmetry between increasing and decreasing fear expressions (T(11) = 15.43, p < 0.001). Last, we tested whether the overall amount of asymmetry differed between increasing and decreasing fear movies. This was not the case, i.e., both movie categories were matched in asymmetry (T (11) = −0.45, p = 0.66).

In the following, we describe subjective ratings obtained using these movie stimuli. Differences for the factor timeline can be attributed to the above observed timeline asymmetries, since static content, overall motion content, and low-level properties were matched across movies and balanced in the factorial design. Subjective perceptual ratings were obtained from human observers about the convincingness, artificialness and emotional intensity of natural facial expression movies. The movies showed dynamically increasing or decreasing facial expressions of fear, either in natural (forward) frame order or in reversed (backward) frame order. While each individual rating has ordinal scales of measurement, statistics were carried out using the mean values from each subject for each condition of each rating, i.e., on continuous values. To test whether the resulting mean values follow Normal distributions, we calculated Shapiro-Wilk- Tests for each dataset. The results confirmed that the mean values do not differ from Normal distribution [rating of intensity: W(112) = 0.984, p = 0.185; rating of artificialness: W(76) = 0.985, p = 0.522; and rating of convincingness: W(76) = 0.975, p = 0.131]. Accordingly, parametric testing was used for further analysis: two-by-two ANOVAs with the factors “timeline” (levels: natural, reversed) and “emotion-direction” (levels: increase, decrease) were calculated for each of the rated features.

The following results were observed. Ratings of emotional intensity (Figure 3A): the ANOVA revealed a main effect of “emotion-direction” [F(1,24) = 71.55, p < 0.001] and a weaker, yet significant main effect of “timeline”

[F(1,24) = 17.56, p < 0.001]. There was no interaction [F(1,24) = 1.7, p = 0.194]. Participants rated emotional intensity higher for increasing compared to decreasing fear, and forward played movies were rated more fearful than reversed movies.

Ratings on artificialness (Figure 3B) showed that decreasing fear stimuli were perceived significantly more artificial than increasing fear, and reversed played movies were perceived more artificial than forward played movies [ANOVA: main effect “emotion-direction” F(1,14) = 10.25, p = 0.005; main effect “timeline” F(1,14) = 12.58, p = 0.002]. There was no interaction [F(1,14) = 0.48, p = 0.497].

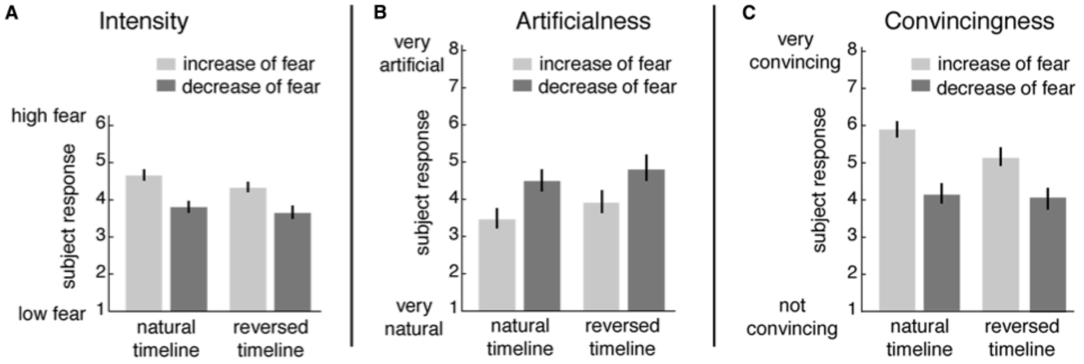


FIGURE 3 | Behavioral ratings (mean and standard error) of dynamic face stimuli. (A) Intensity of fear was rated on a scale from 1 to 6 ranging from “1 = low fear” to “6 = high fear,” (B) artificialness and (C) convincingness were rated on a scale from 1 (“very artificial” or “not convincing”) to 8 (“very natural” or “very convincing”).

Similar to the above, ratings of emotional convincingness (Figure 3C) revealed that decreasing fear was perceived as less convincing than increasing fear [ANOVA: main effect “emotion- direction”: F(1,14) = 49.64, p = 0.001]. Forward played movies were more convincing than reversed movies [ANOVA: main effect “timeline” F(1,14) = 16.79, p < 0.001]. However, post hoc- tests following a significant interaction [ANOVA: “interaction” F(1,14) = 39.35, p < 0.001] showed that the timeline effect can only be found for increasing but not decreasing fear [post hoc paired t-tests for the factor “emotion-direction”:

“natural increase vs. natural decrease” T(18) = −7.89, $p < 0.001$; “artificial increase vs. artificial decrease” T(18) = −5.60, $p < 0.001$; post hoc paired t-tests for the factor “timeline”: “natural increase vs. artificial increase” T(18) = −6.02, $p < 0.001$; “natural decrease vs. artificial decrease” T(18) = −0.81, p = 0.426].

4.5 Discussion

In the current study we examined whether and how time-reversals of dynamic facial expression movies affect subjective experience of human observers. Implicitly, our study also tested whether dynamic facial expressions contain temporal asymmetries, and how they affect perception: if the timeline trajectory of facial action units or of facial muscle movements during the increase of an emotion is not the exact mirror of the timeline during facial relaxation, playing videos in reversed frame order will lead to atypical facial motion trajectories and differences in behavioral ratings. To test this, we recorded natural dynamic facial expressions of increasing and decreasing fear, and played them either forward in the natural frame order (natural timeline) or reversed from last to first frame (reversed timeline). Our design controlled for visual low-level properties, static visual content, and motion energy across the different factors. We found that all three emotional measures, emotional intensity, artificialness and convincingness, were affected by timeline-reversals as well as by the emotional direction of the facial expressions.

Effects of Timeline

The results on the main effects of timeline show that temporally reversed facial dynamics appeared more artificial, less fearful and, in the case of increasing fear, less convincing to our subjects than the natural timeline counterparts. Previous studies on moving objects (Stone, 1998; Chuang et al., 2005; Schultz et al., 2013; Dobs et al., 2014) already indicated that the visual system is well tuned to temporal statistics. Our results provide direct evidence that human perception is well tuned to the familiar temporal order of muscle movement that occurs during simple emotional fear expressions, and that it detects the fine temporal asymmetry of the sequence of these muscle movements during rise and fall of the expression. One process that may partly account for this could be facial mimicry - the imitation of the facial expression perceived by the viewer. Facial mimicry typically accompanies perception of facial emotions, and it occurs spontaneously and rapidly (Hoffman, 1984; Hatfield et al., 1993). Facial mimicry has been shown to help identifying facial expressions (Niedenthal et al., 2001; Krumhuber and Manstead, 2009; Maringer et al., 2011) and is more pronounced when watching dynamic compared to static expressions (Weyers et al., 2006; Sato and Yoshikawa, 2007b; Sato et al., 2008). As reversed facial movies seem not to follow the usual movement trajectories, facial mimicry might not work properly and may therefore contribute to the perception of the facial expressions as more artificial and less convincing. As natural facial expressions follow non-linear trajectories (Cosker et al., 2010), similar effects have been reported for linear expression morphs, that are also rated as less intense and natural than genuine recordings and that are recognized slower and less accurately (Wallraven et al., 2008; Cosker et al., 2010).

Cunningham and Wallraven (2009) also tested the effects of time-reversal of facial dynamics on the identification of different expressions. They compared recognition rates of movies that were played either in forward or backward frame order. Forward, i.e., natural timelines showed slightly higher recognition rates than reversed timelines, independent of the portrayed expression, emphasizing the importance of temporal direction. However, in contrast to us, they did not control for the change in emotion direction that results from timeline reversal. A study of Hill and Johnston (2001) provided another interesting approach that underlined the impact of facial motion. They animated a standard head with various movement patterns of different people. An identity recognition task indicated that subjects could discriminate the correct identities from only the facial dynamics above chance.

Taken together, those findings prove the importance of facial motion on face perception. Our findings further show that humans are sensitive to effects of temporal direction, also when high-level effects of emotion-direction reversal, as well as low- level and static effects are balanced, and provide an account for enhanced emotion recognition for forward movies provided in previous studies.

Effects of Emotion-direction

The results on the main effects of emotion-direction show that decreasing fear was rated as less fearful, more artificial and less convincing than increasing fear.

Fear is a very salient stimulus that draws attention as it signals potential danger. Most studies using dynamic stimuli have used frames from neutral to peak expressions (Kessler et al., 2011; Foley et al., 2012), the full course (i.e., increase followed by decrease; Kilts et al., 2003) or morphs (Kamachi et al., 2001; Labar et al., 2003; Biele and Grabowska, 2006; Sato and Aoki, 2006). Morphs do not contain temporal asymmetries, such that decrease equals reversed increase. To our knowledge, no study has explicitly examined genuine isolated relaxation of fear in detail. Sato et al. (2010) used morphs showing increasing and decreasing fear and found lower intensity ratings for decreasing fear, which we confirm here. In addition, we extend their findings to genuine recordings and show effects that reveal timeline asymmetries in genuine recordings.

What possible reasons could have contributed to the less convincing and more artificial appearance of fear decrease? We can offer two potential explanations for these results. First, it is conceivable that it was easier for our actors to act (or imagine getting into a state of) fear, compared to the reverse that is typically less frequently asked of them to do. Second, the difference may lie in the beholders eye: even though in daily life we may be exposed to just parts of a facial expressions (i.e., when turning to someone), we might be more familiar with increasing than decreasing fear as this has a higher salience for us. Observers may be more used to paying attention to increases of emotional expression, making it more likely to rate what they typically pay less attention to as less convincing. The same account may also explain that when playing the movies in reverse, artificially increasing fear is rated less convincing than natural increasing fear while artificially decreasing fear does not seem to differ much from natural decreasing fear.

Generalization

The question can be raised whether the effects found in our experiment also generalize to other emotional expressions. As mentioned above, Cunningham and Wallraven (2009) did not find any differences between the expressions they tested. However, they point out that some expressions rely more on motion than others. The effects observed in our study clearly depend on the presence of temporal asymmetries in rise and relaxation of the emotional expression. The extent of such asymmetries might vary between different expressions, but this has so far not been systematically quantified in objective or psychometric ways.

The fact that we found relatively robust effects even for one basic emotion expression shows that temporal asymmetries are an important component of facial expressions. This would be of high relevance for both, behavioral as well as neuroimaging studies, as perception and processing are influenced by temporal asymmetries. In particular, asymmetries can by definition only occur in natural movies of facial expressions, but are absent in artificially created linear morphs between two expressions.

There are different theories that try to describe the unfolding of emotional faces. One group comprises discrete-emotion theories (Tomkins, 1982; Ekman, 1992) that define a few basic emotions, each of them coupled to a defined neuromotor program. Once triggered, the expression unfolds completely. However, those theories posit that the action units involved occur in a simultaneous fashion, with similar trajectories and coordinated apexes. In contrast, appraisal theories (Smith and Ellsworth, 1985; Smith, 1989; Ortony and Turner, 1990; Frijda and Tcherkassof, 1997; Smith and Scott, 1997), such as the component-process model of Scherer and Ekman (1984) suggest sequential onsets of action units that can still be modulated during unfolding depending on the situation. Wehrle et al. (2000) aimed to test both types of models and used dynamic stimuli where action units were activated either sequentially or simultaneously but could not find a priority effect for any of them. Our results likewise do not support one theoretical account over the other, but suggest the existence of prototypical unfoldings of emotional expressions that contain temporal asymmetries, at least for fear.

We quantified motion intensity in the facial expression videos, and show significant temporal asymmetries, both within natural recordings of a given emotion-direction, as well across, e.g., natural increasing fear vs. reversed decreasing fear. Our quantification method, however, does not reveal which parts of the faces contains the most asymmetries, and whether asymmetries also exist between rather than within single action units. This interesting question is beyond the scope of the current study and would require analysis of the movement of facial action units based on more detailed facial motion data, e.g., derived using face marker tracking that is not available here.

The importance of facial motion is also evident in neurological or psychiatric disorders. Patients with brain damage, prosopagnostics (Humphreys et al., 1993), the blind (de Gelder et al., 1999), or patients with developmental disorders like autism have been shown to benefit from facial motion in the recognition of emotional expressions and of identity, while failing with static images (Harwood et al., 1999; Back et al., 2007). Mechanisms involved in processing of facial motion trajectories appear to play an important role for these patient groups, making it worthwhile to characterize which facial motion features contribute to the dynamic face advantage in both healthy and patient populations.

Acknowledgments

This work was supported by the Centre for Integrative Neuroscience, University of Tübingen, through the German Excellence Initiative (EXC307) and by the Max Planck Society, Germany.

4.6 References

Ambadar, Z., Schooler, J. W., and Cohn, J. F. (2005). Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. Psychol. Sci. 16, 403–410.

Back, E., Mitchell, P., and Ropar, D. (2007). Do the eyes have it? Inferring mental states from animated facial expressions in children and adolescents with autism. Child Dev. 78, 397–411.

Bartels, A., Zeki, S., and Logothetis, N. K. (2008). Natural vision reveals regional specialization to local motion and to contrast-invariant, global flow in the human brain. Cereb. Cortex 18, 705–717.

Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. J. Exp. Psychol. Hum. Percept. Perform. 4, 373–379.

Biele, C., and Grabowska, A. (2006). Sex differences in perception of emotion intensity in dynamic and static facial expressions. Exp. Brain Res. 171, 1–6.

Bülthoff, H. H., Little, J., and Poggio, T. (1989). A parallel algorithm for real-time computation of optical flow. Nature 337, 549–555.

Chuang, L. L., Vuong, Q. C., Thornton, I. M., and Bülthoff, H. H. (2005). Recognizing novel deforming objects. Front. Comput. Neurosci. 22:158–158.

Cosker, D., Krumhuber, E., and Hilton, A. (2010). "Perception of linear and non- linear motion properties using a FACS validated 3D facial model," Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization. New York, NY: ACM, 101–108.

Cunningham, D. W., and Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. J. Vis. 9, 7–17. doi: 10.1167/9.13.7

de Gelder, B., Vroomen, J., Pourtois, G., and Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. Neuroreport 10, 3759–3763.

Dobs, K., Bülthoff, I., Breidt, M., Vuong, Q. C., Curio, C., and Schultz, J. (2014). Quantifying human sensitivity to spatio-temporal information in dynamic faces. Vision Res. 100, 78–87.

Edwards, K. (1998). The face of time: temporal cues in facial expressions of emotion. Psychol. Sci. 9, 270–276. doi: 10.1111/1467-9280.00054 Ekman, P. (1992). An argument for basic emotions. Cogn. Emot. 6, 169–200.

Foley, E., Rippon, G., Thai, N. J., Longe, O., and Senior, C. (2012). Dynamicfacial expressions evoke distinct activation in the face perception network: a connectivity analysis study. J. Cogn. Neurosci. 24, 507–520.

Fox, C. J., Iaria, G., and Barton, J. J. S. (2009). Defining the face processing network: optimization of the functional localizer in fMRI. Hum. Brain Mapp. 30, 1637–1651.

Freyd, J. J., and Finke, R. A. (1984). Representational momentum. J. Exp. Psychol. 10, 126–132.

Frijda, N. H., and Tcherkassof, A. (1997). "Facial expressions as modes of action readiness," in The Psychology of Facial Expression, eds J. A. Russel and J.M. Fernández-Dols (Cambridge: Cambridge University Press), 78–102.

Finke, R. A., and Freyd, J. J. (1985). Transforming of visual memory induced by implied motions of pattern elements. J. Exp. Psychol. 11, 780–794.

Furl, N., van Rijsbergen, N. J., Kiebel, S. J., Friston, K. J., Treves, A., and Dolan, R. J. (2010). Modulation of perception and brain activity by predictable trajectories of facial expressions. Cereb. Cortex 20, 694–703.

Giese, M. A., and Poggio, T. (2003). Cognitive neuroscience: neural mechanisms for the recognition of biological movements. Nat. Rev. Neurosci. 4, 179–192.

Harwood, N. K., Hall, L. J., and Shinkfield, A. J. (1999). Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. Am. J. Ment. Retard. 104, 270–278.

Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1993). Emotional contagion. Curr. Dir. Psychol. Sci. 2, 96–99.

Hill, H., and Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. Curr. Biol. 11, 880–885. doi: 10.1016/S0960-9822-3

Hoffman, M. L. (1984). "Interaction of affect and cognition in empathy," in Emotions, Cognition, and Behavior, eds C. E. Izard, J. Kagan, and R. B. Zajonc

(Cambridge: Cambridge University Press), 103–131.

Humphreys, G. W., Donnelly, N., and Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces—neuropsychological evidence. Neuropsychologia 31, 173–181.

Jack, R. E., Garrod, O. G., and Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. Curr. Biol. 24, 187–192.

Jellema, T., Pecchinenda, A., and Palumbo, L. (2011). Biases in the perception and affective valence of neutral facial expressions induced by the immediate perceptual history. Vis. Cogn. 19, 616–634.

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., and Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. Perception 30, 875–887.

Kaulard, K., Cunningham, D. W., Bülthoff, H. H., and Wallraven, C. (2012). The MPI facial expression database–a validated database of emotional and conversational facial expressions. PLoS ONE 7:e32321.

Kessler, H., Doyen-Waldecker, C., Hofer, C., Hoffmann, H., Traue, H. C., and Abler, B. (2011). Neural correlates of the perception of dynamic versus static facial expressions of emotion. Psychosoc. Med. 8, Doc03.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., and Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. Neuroimage 18, 156–168.

Knappmeyer, B., Thornton, I. M., and Bülthoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. Vision Res. 43, 1921–1936.

Krumhuber, E. G., and Kappas, A. (2005). Moving smiles: the role of dynamic components for the perception of the genuineness of smiles. J. Nonverb. Behav. 29, 3–24. doi: 10.1007/s10919-004-0887-x

Krumhuber, E. G., and Manstead, A. S. R. (2009). Can Duchenne smiles be feigned? New evidence on felt and false smiles. Emotion 9, 807–820.

Labar, K. S., Crupain, M. J., Voyvodic, J. T., and Mccarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. Cereb. Cortex 13, 1023–1033.

Lander, K., and Bruce, V. (2004). Repetition priming from moving faces.

Mem. Cogn. 32, 640–647.

Lander, K., Chuang, L., and Wickham, L. (2006). Recognizing face identity from natural and morphed smiles. Q. J. Exp. Psychol. 59, 801–808.

Lederman, S. J., Kilgour, A., Kitada, R., Klatzky, R. L., and Hamilton, C. (2007). Haptic face processing. Can. J. Exp. Psychol. 61, 230–241.

Marian, D. E., and Shimamura, A. P. (2013). Contextual influence on dynamic expressions. Am. J. Psychol. 126, 53–65.

Maringer, M., Krumhuber, E. G., Fischer, A. H., and Niedenthal, P. M. (2011). Beyond smile dynamics: mimicry and beliefs in judgments of smiles. Emotion 11, 181–187.

Niedenthal, P. M., Brauer, M., Halberstadt, J. B., and Innes-Ker, Å. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. Cogn. Emot. 15, 853–864.

Ortony, A., and Turner, T. J. (1990). Whats basic about basic emotions. Psychol. Rev. 97, 315–331.

O'Toole, A. J., Roark, D. A., and Abdi, H. (2002). Recognizing moving faces: a psychological and neural synthesis. Trends Cogn. Sci. 6, 261–266.

Pilz, K. S., Thornton, I. M., and Bülthoff, H. H. (2005). A search advantage for faces learned in motion. Exp. Brain Res. 171, 436–447.

Pollick, F. E., Hill, H., Calder, A., and Paterson, H. (2003). Recognising facial expression from spatially and temporally modified movements. Perception 32, 813–826.

Reinl, M., and Bartels, A. (2014). Face processing regions are sensitive to distinct aspects of temporal sequence in facial dynamics. Neuroimage 102, 407–415.

Richoz, A. R., Jack, R. E., Garrod, O. G., Schyns, P. G., and Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. Cortex 65, 50–64.

Sato, W., and Aoki, S. (2006). Right hemispheric dominance in processing of unconscious negative emotion. Brain Cogn. 62, 261–266.

Sato, W., Fujimura, T., and Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial expressions. Int. J. Psychophysiol. 70, 70–74.

Sato, W., Kochiyama, T., and Yoshikawa, S. (2010). Amygdala activity in response to forward versus backward dynamic facial expressions. Brain Res. 1315, 92–99.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., and Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. Brain Res. Cogn. Brain Res. 20, 81–91.

Sato, W., and Yoshikawa, S. (2004). The dynamic aspects of emotional facial expressions. Cogn. Emot. 18, 701–710.

Sato, W., and Yoshikawa, S. (2007a). Enhanced experience of emotional arousal in response to dynamic facial expressions. J. Nonverbal. Behav. 31, 119–135.

Sato, W., and Yoshikawa, S. (2007b). Spontaneous facial mimicry in response to dynamic facial expressions. Cognition 104, 1–18.

Scherer, K. R., and Ekman, P. (1984). Approaches to Emotion. Hillsdale, NJ: Psychology Press.

Schultz, J., Brockhaus, M., Bulthoff, H. H., and Pilz, K. S. (2013). What the human brain likes about facial motion. Cereb. Cortex 23, 1167–1178.

Schultz, J., and Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. Exp. Brain Res. 194, 465–475.

Simons, R. F., Detenber, B. H., Roedema, T. M., and Reiss, J. E. (1999). Emotion processing in three systems: the medium and the message. Psychophysiology 36, 619–627.

Smith, C. A. (1989). Dimensions of appraisal and physiological response in emotion. J. Pers. Soc. Psychol. 56, 339–353.

Smith, C. A., and Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. J. Pers. Soc. Psychol. 48, 813–838.

Smith, C., and Scott, H. (1997). “A componential approach to the meaning of facial expressions,” in The Psychology of Facial Expression, eds J. A. Russel and J. M. Fernández-Dols (Cambridge: Cambridge University Press), 229–254.

Stone, J. V. (1998). Object recognition using spatiotemporal signatures. Vision Res. 38, 947–951.

Thornton, I. M., and Kourtzi, Z. (2002). A matching advantage for dynamic human faces. Perception 31, 113–132.

Tomkins, S. (1982). “Affect theory,” in Emotion in the Human Face, 2nd Edn, ed P. Ekman (Cambridge: Cambridge University Press), 353–397.

Trautmann, S. A., Fehr, T., and Herrmann, M. (2009). Emotions in motion: dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. Brain Res. 1284, 100–115.

Wallraven, C., Breidt, M., Cunningham, D. W., and Buelthoff, H. H. (2008). Evaluating the perceptual realism of animated facial expressions. ACM Trans. Appl. Percept. 4, 1–20.

Wehrle, T., Kaiser, S., Schmidt, S., and Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. J. Pers. Soc. Psychol. 78, 105–119.

Weyers, P., Mühlberger, A., Hefele, C., and Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. Psychophysiology 43, 450–453.

Yoshikawa, S., and Sato, W. (2008). Dynamic facial expressions of emotion induce represental momentum. Cogn. Affect. Behav. Neurosci 8, 25–31.

5 The impact of temporal asymmetries on the evaluation of dynamic facial expressions in Autism Spectrum Disorder

5.1 Abstract

People with Autism Spectrum Disorder (ASD) show impairments in processing faces. However, it is unclear, if these impairments are mainly for static facial cues or also extend to dynamic ones. Using a previously established design, we investigated the sensitivity of ASD and typical developed (TD) participants towards manipulations of temporal information in facial expressions. We presented genuine movie recordings of five basic facial expressions in a 2x2 factorial design. Movies of increasing or decreasing expression intensity (factor "emotion-direction") were played either in forward (natural) or reversed frame order (factor "timeline"). The latter manipulation results in altered temporal movement trajectories of the facial muscles. This design provides full control over static and dynamic features while exclusively manipulating the temporal trajectory of facial motion. We obtained expression and timeline recognition rates as well as intensity and authenticity ratings.

5.2 Introduction

For the development of social skills, one fundamental aspect seems to be an inherent interest in faces. Normally, already newborns preferentially turn to faces (Carey, 1981; Ellis 1990) but this preference is missing in children later diagnosed with Autism. Often, general face processing deficits are described for children and adults diagnosed with ASD (Davies et al., 1994; Gepner et al., 2001) but findings are inconsistent. Deficits have primarily been found in affect-matching paradigms (Davies et al., 1994; Piggot et al., 2004; Rump et al., 2009) and less in tasks demanding explicit labelling of expressions (Kätsyri and Sams, 2008; Piggot et al., 2004; Rutherford and Towns, 2008). Some studies also report longer reaction times in ASD despite similar recognition accuracy (Piggot et al., 2004; Corbett et al., 2009; Kleinhans et al., 2009). This could be indicative of fundamental difficulties in facial perception in ASD, whereas deficits in accuracy in some patients can be compensated for by coping strategies at the cost of reaction time. Other potential reasons for the discrepant findings might be the heterogeneity of the experimental operationalisation as well as the diverge characteristics of the ASD groups.

The above-mentioned studies have used photographs of faces as stimuli. In real life, however, we usually encounter moving faces which we percieve and process effortlessly. In TD participants, such dynamic stimuli have been shown to enhance emotion recognition (Harwood et al., 1999; Wehrle et al., 2000; Ambadar et al., 2005; Biele and Grabowska, 2006; Cunningham and Wallraven, 2009; Weyers et al., 2006). Subjects are further able to reproduce the sequence of movement of facial expressions from photographs (Edwards, 1998) and can detect even small variations in facial movement trajectories (Dobs et al., 2014). More detailed analysis on the evolvement of facial expressions have shown that facial expressions have asymmetric motion trajectories between increasing and decreasing intensity (Curio et al., 2006; Dobs et al., 2014; Jack et al., 2014). In two previous studies we showed that our perception and neuronal processing is influenced by those asymmetries (Reinl and Bartels, 2015; 2014). In a well-controlled design, we presented movies of increasing and decreasing fear in either forward or reversed frame order. Reversing the frame order results in altered motion trajectories of the facial expression. We found that ratings of the artificialness of facial movement, the convincingness of expression portrayal, and the intensity of the expressions were influenced by frame order reversal and emotion direction (Reinl and Bartels, 2015). Additionally, we could show that brain areas dedicated to face processing were differentially affected by our manipulation (Reinl and Bartels, 2014). Unfortunately, the use of dynamic face stimuli in autism research has not shed light on the controversial findings for their face processing abilities. Both, Loveland et al., (1997) and Gepner et al., (2001) find no differences between ASD and TD children and adults using dynamic face stimuli. However, Loveland et al., (1997) did not include a static control condition and Gepner et al., (2001) already detects no group differences for static faces. In contrast, Lindner and Rosén, (2006) report lower performance for both static and dynamic facial expressions in ASD children. Interestingly, ASD children were able to re-create the movement sequence of some facial expressions and performed even better than TD for spoken words (Grossman and Tager-Flusberg, 2008). Emotion recognition abilities however might be influenced by the speed of stimulus presentation. When movies were slowed down in pace, Tardif et al., (2007) found better performance in ASD participants than for stimuli presented at normal speed. This effects however was not found by Bal et al., (2010) who still reports longer reaction times and some impairments in expression recognition despite using videos of faces that change from neutral to expressional appearance very slowly. It is therefore unclear wether ASD is associated to systematic impairments of dynamic face perception and if so, which aspect of facial dynamics lead to such impairments.

In this study, we hence aim to investigate the sensitivity of participants with ASD towards fine variations in facial movement. We presented normal paced, genuine recordings of five facial expressions (fear, happy, surprise, anger and disgust). We used our previously established two-factorial design (Reinl and Bartels, 2014; 2015) with movies comprising increasing and decreasing expression intensity (factor "emotion-direction") that were shown in either natural or reversed frame order (factor "timeline"). This approach matched all visual aspects aside from directionality of motion trajectories. It therefore enabled us to investigate the effects that timeline reversal as well as emotional directionality have on the perception and evaluation of the facial stimulus. To test the effects of our manipulation, we obtained expression recognition rates and ratings about intensity and authenticity of the presented expressions. To further test whether possible timeline effects are due to the explicit knowledge of our manipulation, or rather due to implicit processing by the participants, we also asked subjects to estimate which timeline present in each stimulus.

5.3 Methods

Participants

22 adults (12 male, mean age: 37 ± 9) diagnosed with ASD by clinical interview and 22 TD individuals (12 male, mean age: 37 ± 9) with no self-reported psychiatric or neurological history were recruited for this study. ADS and control group were individually matched according to gender, age and educational level.

Stimuli and Procedure

Stimulus generation and conditions.

Stimuli included colored short movie clips of four Caucasian actors portraying fearful, surprised, happy, angry and disgusted facial expressions. Movies were recorded prior to the experiment. Actors currently working in a local theater were asked to imagine a situation to portray the above-mentioned emotional expressions. They were asked to start with a neutral face, go to peak expression and relax back to a neutral condition. During that they were told to keep their head still to minimize rigid head movements. Three repetitions were recorded from each actor and one movie of each expression from each actor was afterwards selected by visual inspection. Requirements were a recognizable expression with a clear increase and decrease, without excessive eye blinking and no or only little head movement. Movies were then cut into sections containing either increasing or decreasing expressions, ranging from low to high intensity or vica versa using VirtualDub (virtualdub.org). All movie clips were presented in the original as well as reversed frame order, resulting in four conditions: increasing and decreasing facial expression in original frame order (natural timeline), and decreasing and increasing facial expression in reversed frame order (reversed timeline). The experimental design is illustrated in Figure 1.

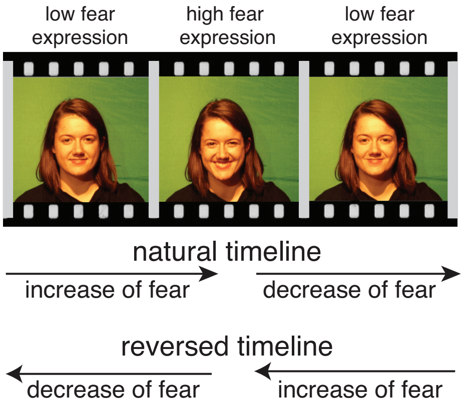


Fig. 1: Experimental paradigm: videos of facial expressions were shown in four conditions: increase and decrease of expression (factor emotion-direction) in natural as well as reversed frame order (factor timeline).

Stimulus presentation.

Participants were placed in front of a computer monitor and performed different rating tasks on the stimuli. Movies were presented in a randomized sequence that was counterbalanced across subjects. During the experiment, each trial started with a fixation cross in the center of a gray screen presented for 1200 ms. The movie was shown with normal speed, but the first and last frames were additionally presented before and after the movie for 150 ms and 100 ms, respectively. This was done to enhance the recognizability of the movies by reducing forward- or backward masking effects of the gray screen. This resulted in a mean duration of 850ms of the movies. After each movie, another grey screen was presented for 800 ms before the rating instructions appeared. Subjects had no time limit to answer. After giving the answer, the screen turned gray again for another 800 ms before the next rating screen appeared (see figure 2).

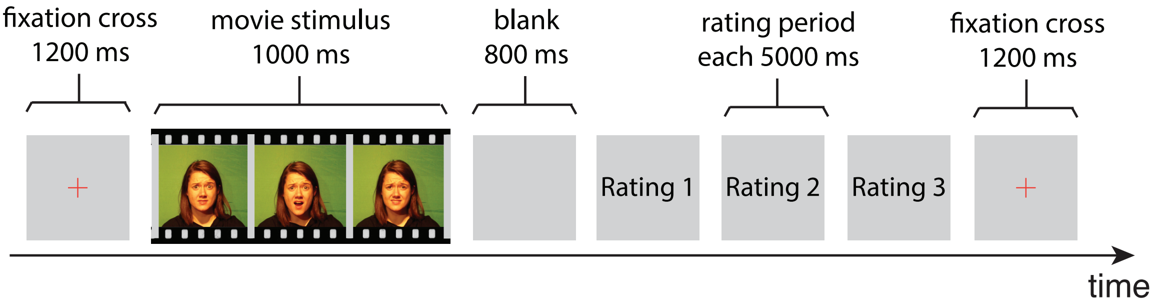


Fig. 2: Illustration of trial sequence.

Naïve and informed rating sessions

The study consisted of two sessions, in each of which the participants viewed the full set of movie stimuli (yet in distinct order). In session one, participants were kept naive with regards to the stimulus manipulation. Prior to session two, participants were informed about the stimulus manipulation. For all rating tasks, participants had no time limitation for their answer. Before each session, participants performed a short test run of 10 stimuli to get used to the task and the movie presentation.

In session one, participants were asked to respond by button press to the following rating questions: 1. Which expression was shown in the movie? (select correct expression from five option presented on screen), 2. How certain are you about your answer? (select from 1- very uncertain to 5 – very certain), 3. How authentic was the expression presented (select from 1 – not authentic to 7 – very authentic). Authenticity described if the actor was actually feeling the presented emotion.

The second session then included the following questions: 1. How intense was the emotion? (select from 1 - very weak to 7 - very strong), 2. Was the movie played forward or backward? (select either 1 - forward or 2 - backward), 2. How certain are you about your answer? (select from 1- very uncertain to 5- very certain).

Data analysis

Data was analyzed using Matlab 2014b and SPSS 22 (IBM Statistic). For each rating type, we calculated mixed design ANOVAs with three factors: two within-subject factors “timeline” (natural versus reversed) and “emotion-direction” (increase versus decrease) and one between-subject factor “group” (ASD versus TD). Bonferroni-Holm-correction was applied for the number of ANOVAs. Additionally, dependent t-tests were calculated to compare the mean recognition rates for positive versus negative expressions. Bonferoni-correction was applied for the number of t-tests, such that p-values below 0.025 were considered as significant.

5.4 Results

We calculated mixed 3-way ANOVAs with two within-subject factors “timeline” (natural versus reversed) and “emotion-direction” (increase versus decrease) and one between-subject factor “group” (ASD versus TD).

Correct recognition of the expression

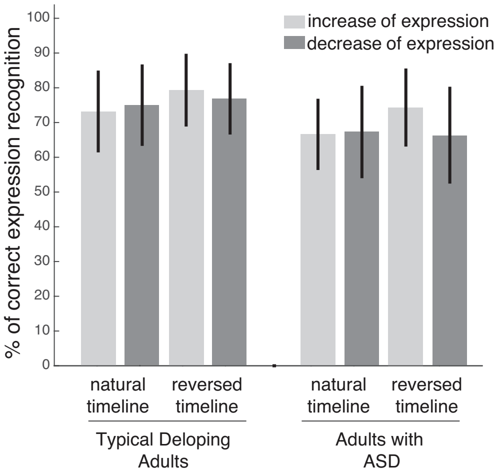


Fig. 3: Percentage of correct expression recognition. Error bars represent standard deviation.

First, we evaluated accuracy of expression recognition (Fig.3). Accuracy is given as the percentage of correct identification (Fig.3). The 3-way ANOVA revealed a main effect of the participant group (F(1,42) = 8.01, p = 0.007, ŋ2 = 0.16) with lower recognition rates in the ASD group. We further obtained a main effect of timeline (F(1,42) = 9.83, p = 0.003 ŋ2 = 0.19), resulting in higher recognition rates for reversed than natural timelines. The interaction of timeline and emotion-direction (F(1,42) = 4.54, p = 0.038, ŋ2 = 0.1) did not survive Bonferoni-Holm-correction. There was no main effect of emotion-direction or any further interactions.

Since several prior studies have found that the recognition of negative expressions is more impaired than that of positive emotions (Ashwin et al., 2006, Boraston et al., 2007; Howard et al., 2000; Corden et al., 2008; Pelphrey et al., 2007), we wanted to test whether the same was true in the current data. We split the data into negative (fear, disgust,anger) and positive (surprise, happy) emotions, averaging over timeline and emotion-direction conditions to test whether the performance of ASD and TD differed depending on the valence using paired t-test. Negative emotions were significantly better recognized by TD (t(21) = -3.35, p = 0.003, performance: ASD participants

66%, TD participants 76%), while there was no difference for positive emotions (t(21) = -1.17, p = 0.256, performance: ASD participants 72%, TD participants 76%).

Confidence rating for expression recognition

Second, we tested whether the confidence of the expression recognition were influenced by any factor of the 2x2 stimulus design (Fig.4). The 3-way ANOVA revealed a main effect for participant group (F(1,42) = 20.53, p = 0.000, ŋ2 = 0.33) with higher confidence in TD compared to ASD. We further obtained a main effect of emotion-direction (F(1,42) = 43.18, p = 0.000, ŋ2 = 0.51), resulting in higher confidence for increasing than decreasing emotions. There was no effect for timeline, nor any interactions.

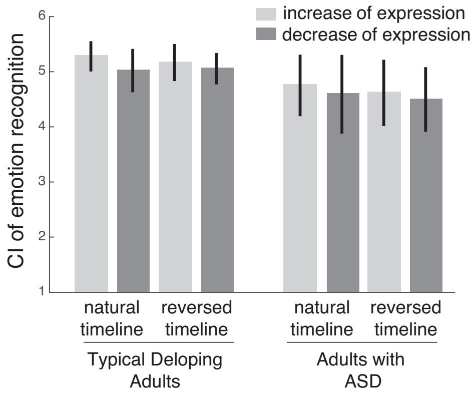


Fig. 4: Confidence rating for expression recognition. Error bars represent standard deviation

Rating of Authenticity

We aquired ratings of authenticity, which described if the actor was actually feeling the emotion that he or she presented (Fig.5). We obtain a main effect of group (F(1,42) = 11.16, p = 0.002, ŋ2 = 0.21) with higher authenticity ratings for TD compared to ASD as well as a main effect of emotion-direction (F(1,42) = 17.73, p = 0.000, ŋ2 = 0.3) with higher authenticity ratings for increasing than decreasing emotions. There was no effect for timeline, nor any interactions.

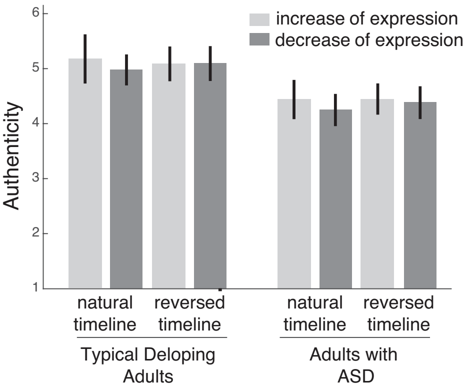


Fig. 5: Rating of Authenticity. Error bars represent standard deviation.

Intensity of emotions

For ratings of perceived emotion intensity (Fig.6), we obtained a main effect of emotion-direction (F(1,42) = 7.99, p = 0.007, ŋ2 = 0.16), with in higher intensity ratings for increasing than decreasing emotions. There was an interaction between timeline and emotion direction (F(1,42) = 62.88, p = 0.000, ŋ2 = 0. 06), showing higher intensity ratings for increasing emotions only for the natural timeline. A main effect of group did not survive Bonferoni-Holm-correction (F(1,42) = 5.44, p = 0.025, ŋ2 = 0.12) and there was no effect for for timeline or any further interactions.

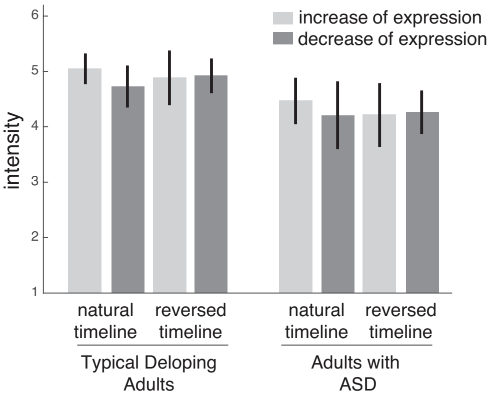


Fig. 6: Rating of the emotion intensity. Error bars represent standard deviation.

Correct identification of timeline

We asked our subjects to decide on the timeline of each stimulus, i.e. if it was played in original or reversed frame order. Responses are given as percentage of correct identification (Fig.7). The 3-way ANOVA revealed a main effect of timeline with higher recognition rates for natural timeline (F(1,42) = 12.04, p = 0.001, ŋ2 = 0.22). A significant interaction between timeline and emotion-direction showed however, that this effect was only found for increasing expressions (F(1,42) = 130.64, p = 0.000, ŋ2 = 0.76). A main effect of emotion direction (F(1,42) = 6.67, p = 0.013, ŋ2 = 0.13) did not survive Bonferoni-Holm-correction, neither

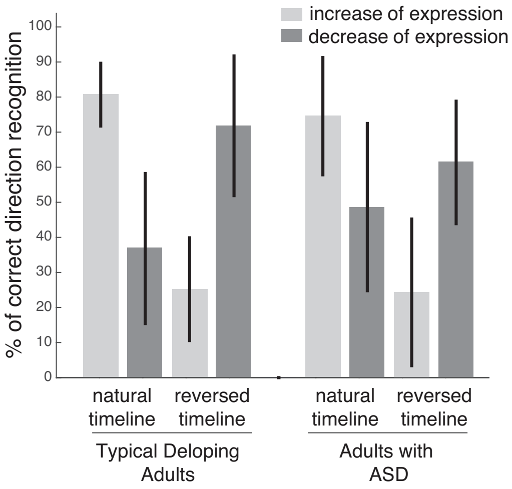


Fig. 7: Percentage of correct timeline recognition. Error bars represent standard deviation.

did a tripple interaction of all factors (F(1,42) = 4.2, p = 0.047, ŋ2 = 0.09). There was no main effect of participant group or any further interaction.

Confidence rating for identification of timeline

Subjects were asked to further indicate the certainty of their decision on the timeline recognition (Fig.7). The 3-way ANOVA revealed a main effect of emotion-direction (F(1,42) = 15.15, p = 0.000, ŋ2 = 0.27) with subjects being more certain about their timeline identification for increasing expression direction. However, this effect is limited to natural timelines, indicated by a significant interaction of timeline and expression-direction (F(1,42) = 39.44, p = 0.000, ŋ2 = 48). Again, the tripple interaction between all factors did not survive Bonferroni-Holm correction (F(1,42) = 6.13, p = 0.016, ŋ2 = 0.13). There was no significant group effect or any further interactions.

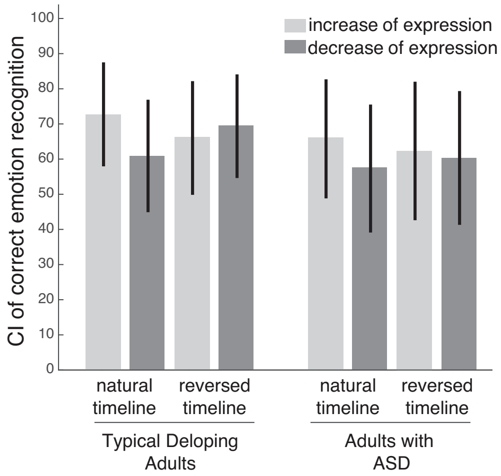


Fig. 8: Confidence rating for expression recognition. Error bars represent standard deviation

5.5 Discussion

In our study we examined whether participants with ASD were differentially affected by fine differences in temporal trajectories of dynamic facial expressions than TD controls. For this purpose, we used a 2x2 factorial stimulus design that fully controlled for low-level stimulus aspects while altering the dynamics of facial muscle trajectories. We presented movies showing expressions with increasing and decreasing intensity (factor „emotion-direction“), that were played in the natural as well as reversed frame order (factor „timeline“). We investigated recognition rates of expression type and timeline and obtained ratings of intensity and authenticity of the expressions.

Manipulation of motion trajectory

A main effect of timeline shows higher recognition rates for the identification of the natural frame order. Additionally we find a significant interaction between “timeline” and “emotion-direction”. Movies of naturally increasing expressions and their timeline-reversed counterparts were better recognized than movies of naturally decreasing expressions and their corresponding timeline-reversed variants. A tripple interaction with the group factor did not survive Bonferoni-Correction and their was no main effect of group. We also do not find interaction of timeline and group for the other ratings. ASD participants therefore seem not to be differentially affected by our timeline manipulation than the control group.

According to our results, we are sometimes able to identity the familiar temporal order of muscle movement and detect violations from that. The underlying process that might at least in part account for that could be facial mimicry. Facial mimicry describes automatic facial reactions in response to expressions that we see in other people – a mechanism thought to enhance social interaction, mediated by the so-called mirror neuron system (MNS). However, in daily life, we seldom see an expression fully evolving and declining but often see just parts of it, i.e. when we turn towards someone. As increasing expressions are usually of higher salience to us and might therefore catch our attention more often as decreasing expressions, we might be more familiar with them and better tuned towards their temporal properties.

Early studies on imitation abilities in ASD children have provided evidence compatible with a dysfunction of the mirror neuron system (Hobson and Lee, 1999; Ohta, 1987; Rogers et al., 1996; Rogers et al., 2003). A study of Scambler (Scambler et al., 2007) investigated emotional responsiveness to emotional displays and found reduced mimicry of two year old children with ASD compared to TD children. It is unclear though, if at this early age, such facial reactions are really accomplished by the MNS or rather caused by other processes like arousal, exploration, and emotional contagion (see Beall et al., 2008). Many studies report comparable performance of ASD and TD children on goal-directed tasks (Aldridge et al., 2000; Carpenter et al., 2001; Hamilton et al., 2007), imitation of object-directed actions (Ingersoll et al., 2003), normal imitation of object-use (Stone et al., 1997) and other imitation tasks including meaningful actions (Rogers et al., 1996; Brown and Whiten, 2009). In contrast, ASD children have been described to be impaired in theory of mind tasks (Hamilton et al., 2007; Happé, 1995), imitation of meaningless actions or gestures (Rogers et al., 1996; Stone et al., 1997; Ohta, 1987; Charman et al., 1997) and tasks involving emotional facial expressions (Dapretto et al., 2006). As the MNS is a part of the motor system, it operates on both, goal-directed tasks and facial mimicry. Therefore, the theory of a complete malfunctioning of the MNS in ASD has been questioned. Furthermore, when McIntosh and Decker (2006) explicitly instructed adult participants with ASD to mimic facial expressions, they were able to perform but they did not spontaneously mimic the expression without instruction. Similar findings were also reported by other studies (Beall et al., 2008; Ingersoll, 2008; McDuffie et al., 2007). In contrast, Rogers et. al., (2003) report deficit in voluntary imitation of some facial actions performed by a live model. However, they used observers to collect and evaluate responses, which might be less sensitive than using Electromyography. Additionally, people with ASD have been reported to exhibit flattened or blended expressions (Yirmiya et al., 1989), which may have made it difficult to observe facial mimicry in them.

Based on this findings, we would expect our ASD participants to be differentially affected by our timeline manipulation. However, this was not the case. One potential reason for that might be the fact that we did not tested recognition rates seperatly for the different expressions. The findings for emotion recognition however indicate difficulties that at least partially depend on the type of expression. Analyzing the recognition rates seperatly for each expression type could be more sensitive towards potential differences between the ASD and TD group.

Emotion recognition

A main effect of "timeline" was also found for expression recognition, showing surprisingly lower recognition rates for natural compared to reversed expressions. This is in contrast to a previous study by Cunningham and Wallraven (2009), who report slightly higher recognition rates of facial expression movies presented in forward compared to reversed frame order. However, they did not control for emotion direction, so that reversing the frame order also reversed the emotion-direction. It is therefore unclear, wether the slightly higher performance for expressions presented in forward frame order found by Cunningham and Wallraven (2009) resembles a true timeline effect or rather an emotion effect or the mixture of both.

Additionally, participants with ASD were less accurate in identifying emotional expressions and this effect was mainly driven by negative emotions. This is consistent with previous findings that recognition rates of basic emotions can depend on the valence. The recognition of negative expressions seem to be stronger impaired than that of positive emotions (Ashwin et al., 2006; Boraston et al., 2007; Howard et al., 2000; Corden et al., 2008; Pelphrey et al., 2007). However, this is not universally true. In a sequence recreation task, Grossman and Tager-Flusberg (2008) found best performance of both TD and ASD for fearful expressions, and Loveland et al., (1997) did not find any differences between expressions at all. Furthermore, the detected impairments also depend on the type of task used in the experiments. Affect-matching paradigms appear to more difficult (Davies et al., 1994; Piggot et al., 2004; Rump et al., 2009) than expression labelling tasks (Kätsyri and Sams, 2008; Rutherford and Towns, 2008). Using an expression labelling task, we find effects mainly for negatively valenced expressions. Compared to the majority of studies, however we used genuine recordings of facial expressions instead of face photographs. The findings for emotion recognition abilities continued to be controversial with the use of dynamic face stimuli. Some studies found no differences in recognition rates between ASD and TD paricipants (Loveland et al., 1997; Gepner et al., 2001) while others still report worse performance or slower reaction times in the ASD group (Bal et al., 2010; Lindner and Rosén, 2006). Enticott et al., (2014) compared expression recognition rates for static and dynamic faces, and did not find a consistent facilitation effect for dynamic faces but differential effects depending on the facial expression. Unfortunately, we did not include static expressional faces in our experiment. Gepner and Féron (2009) hypothesized that an impaired integration of visual motion interferes with the processing of dynamic stimuli. They suggest that a reduced pace of stimulus presentation can help ASD participants to improve performance. This effect has been found by Tardif et al., (2007), however Bal et al., (2010) still detects worse performance for ASD participants when using movies with very slow expressional changes. We used normal paced movies and find reduced recognition accurarcy in the ASD group which mainly consisted of participants with Asperger diagnosis. There might be differences for Autism and Asperger Syndrome though, as children with autism exhibit weaker postural reaction to environmental motion, while children with Asperger syndrome show enhanced reactivity (Gepner et al., 1995; Gepner and Mestre, 2002). In our data, ASD participants were also less certain about their decision. However, the relatively low average performance of only 76% of the control group suggests that our stimuli were generally not easily distinguishable, which might have added to the impairments found for ASD participants.

We do not find a main effect of emotion direction, which means that neither TD nor ASD participants differed in their performance for increasing and decreasing expressions. Still, both groups were less confident about their decision when labeling decreasing compared to increasing expressions. Additionally, ASD were less confident about their decision on the presented expression type. This is in line with self-reported insecurity of the ASD participants during the experiment. In a next step, a closer look at reaction times might be interesting to test if the lower confidence of ASD participants is also reflected in prolonged reaction times compared to the control group.

Intensity and authenticity rating

The present data showed that both, ASD and TD groups rated increasing emotions as more intense than decreasing emotions, however only for natural timeline movies. This is consistent with the results of our previous study using a distinct stimulus set and TD participants only (Reinl and Bartels, 2015). However, a major difference is that in our prior study the effect was not restricted to natural timeline movies. Instead, we found a main effect of timeline with lower intensity values for reversed timelines. This discrepancy may be due to potentially different extents of asymmetry between increase and decrease of the different expressions. We showed in our previous study, that fearful expressions are significantly asymmetric between increasing and decreasing intensity. As we combine five different emotions in this study, the differences in asymmetry for the different emotions might account for the discrepancy of the effects. Nevertheless, this findings underline the use of genuine movie recording for the research of face processing.

We found that both groups also rated increasing emotion as more authentic then decreasing emotions, independent from the timeline direction. As already mentioned above, in daily life, we often just see parts of an expression unfolding. The higher salience of increasing expressions might result in higher familiarity with increasing expressions and therefore makes it likely to also rate them as more authentic than decreasing ones.

Interestingly, ASD participants gave generally lower authenticity ratings to all stimuli group compared to TD group. Explicit authenticity ratings have not been investigated in ASD before. Walsh et al., (2014) however showed higher tolerance towards exaggerated schematic facial expressions, presumably appearing less authentic if explicitly evaluated, in ASD subjects. In addition, Adolphs et al., (2001) found that ASD participants gave high trustworthiness ratings for faces that were usually rated as least trustworthy by TD participants. Both studies are in contrast to our findings and indicate difficulties of ASD participants for higher-level jugdement of faces and facial expressions. However, most of the ASD, but none of the TD participants in our study reacted surprised when they were asked to do authenticity ratings. They were aware that all videos were posed by actors for camera and therefore had only limited authenticity. Therefore they questioned the sense of the authenticity rating as non of the actors would indeed feel the emotion. We then explained that actors were told to imaging a situation were they would react with the desired expression and therefore expressions would have some authenticity. Still, this might explain the overall lower authenticity rating of ASD.

Outlook

As only four stimuli per expression category were included in the current design, we have averaged the data across all emotions in our analyses. There might be specific effects for the different expressions, however, due to the limited number of trials for each category, we were not able to split our analysis to investigate the effects of our stimulus manipulation for each expression separately. As the variability of asymmetries of increasing and decreasing intensities might differ among different expression types, averaging might mask emotion-specific timeline effects or timeline - group interactions. It is possible, that expressions with high motion intensity like fear or surprise have higher asymmetries than an expression like anger that are often rather subtle. As a next step, it will be interesting to quantify the amount of motion asymmetry using the same technique as described in the previous study (Reinl and Bartels, 2015). With splitting the data according to high or low motion asymmetry, we could increase the number of trials per catergory and test for further effects.

5.6 References

Adolphs, R., Sears, L., Piven, J., 2001. Abnormal processing of social information from faces in autism. J Cogn Neurosci 13, 232–240.

Aldridge, M.A., Stone, K.R., Sweeney, M.H., Bower, T.G.R., 2000. Preverbal children with autism understand the intentions of others. Developmental science 3, 294–301.

Ambadar, Z., Schooler, J.W., Cohn, J.F., 2005. Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. Psychological science : a journal of the American Psychological Society / APS 16, 403–410.

Ashwin, C., Chapman, E., Colle, L., Baron-Cohen, S., 2006. Impaired recognition of negative basic emotions in autism: a test of the amygdala theory. Soc Neurosci 1, 349–363.

Bal, E., Harden, E., Lamb, D., Van Hecke, A.V., Denver, J.W., Porges, S.W., 2010. Emotion recognition in children with autism spectrum disorders: relations to eye gaze and autonomic state. J Autism Dev Disord 40, 358–370.

Beall, P.M., Moody, E.J., McIntosh, D.N., Hepburn, S.L., Reed, C.L., 2008. Rapid facial reactions to emotional facial expressions in typically developing children and children with autism spectrum disorder. J Exp Child Psychol 101, 206–223.

Biele, C., Grabowska, A., 2006. Sex differences in perception of emotion intensity in dynamic and static facial expressions. Exp Brain Res 171, 1–6.

Boraston, Z.L., Corden, B., Miles, L.K., Skuse, D.H., Blakemore, S.-J., 2007. Brief Report: Perception of Genuine and Posed Smiles by Individuals with Autism. J Autism Dev Disord 38, 574–580.

Brown, J.D.B., Whiten, A., 2009. Elicited imitation in children and adults with autism: is there a deficit? Journal of Intellectual and Developmental Disability 29, 147–163.

Carey, S., 1981. The development of face perception. In G. Davies, H.Elli, & J. Shepard (Eds), Perceiving and remembering faces (pp. 12-24). London: Academic Press.

Carpenter, M., Pennington, B.F., Rogers, S.J., 2001. Understanding of Others' Intentions in Children with Autism. J Autism Dev Disord 31, 589–599.

Charman, T., Swettenham, J., Baron-Cohen, S., 1997. Infants with autism: an investigation of empathy, pretend play, joint attention, and imitation. Developmental Psychology 33, 781-789.

Corbett, B.A., Carmean, V., Ravizza, S., Wendelken, C., Henry, M.L., Carter, C., Rivera, S.M., 2009. A functional and structural study of emotion and face processing in children with autism. Psychiatry Res 173, 196–205.

Corden, B., Chilvers, R., Skuse, D., 2008. Avoidance of emotionally arousing stimuli predicts social-perceptual impairment in Asperger's syndrome. Neuropsychologia 46, 137–147.

Cunningham, D.W., Wallraven, C., 2009. Dynamic information for the recognition of conversational expressions. JOV 9, 7–7.

Curio, C., Breidt, M., Kleiner, M., Vuong, Q.C., Giese, M.A., Bülthoff, H.H., 2006. Semantic 3D motion retargeting for facial animation, the 3rd symposium. ACM, New York, New York, USA.

Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y., Iacoboni, M., 2006. Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. Nat Neurosci 9, 28–30.

Davies, S., Bishop, D., Manstead, A.S., Tantam, D., 1994. Face perception in children with autism and Asperger's syndrome. J Child Psychol Psychiatry 35, 1033–1057.

Dobs, K., Bülthoff, I., Breidt, M., Vuong, Q.C., Curio, C., Schultz, J., 2014. Quantifying human sensitivity to spatio-temporal information in dynamic faces. Vision Res 100, 78–87.

Edwards, K., 1998. The face of time: Temporal cues in facial expressions of emotion. Psychological science : a journal of the American Psychological Society / APS 9, 270–276.

Ellis, H.D. (1990) Developmental trends in face recognition. The Psychologist, 3, 124-129.

Enticott, P.G., Kennedy, H.A., Johnston, P.J., Rinehart, N.J., Tonge, B.J., Taffe, J.R., Fitzgerald, P.B., 2014. Emotion recognition of static and dynamic faces in autism spectrum disorder. PCEM 28, 1110–1118.

Gepner, B., Deruelle, C., Grynfeltt, S., 2001. Motion and emotion: a novel approach to the study of face processing by young autistic children. J Autism Dev Disord 31, 37–45.

Gepner, B., Féron, F., 2009. Autism: a world changing too fast for a mis-wired brain? Neurosci Biobehav Rev 33, 1227–1242.

Gepner, B., Mestre, D., 2002. Rapid visual-motion integration deficit in autism. Trends Cogn Sci (Regul Ed) 6, 455.

Gepner, B., Mestre, D., Masson, G., de Schonen, S., 1995. Postural effects of motion vision in young autistic children. NeuroReport 6, 1211–1214.

Grossman, R.B., Tager-Flusberg, H., 2008. Reading faces for information about words and emotions in adolescents with autism. Research in Autism Spectrum Disorders 2, 681–695.

Hamilton, A.F. de C., Brindley, R.M., Frith, U., 2007. Imitation and action understanding in autistic spectrum disorders: How valid is the hypothesis of a deficit in the mirror neuron system? Neuropsychologia 45, 1859–1868.

Happé, F.G.E., 1995. The Role of Age and Verbal Ability in the Theory of Mind Task Performance of Subjects with Autism. Child Dev 66, 843–855.

Harwood, N.K., Hall, L.J., Shinkfield, A.J., 1999. Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. Am J Ment Retard 104, 270–278.

Hobson, R.P., Lee, A., 1999. Imitation and identification in autism. J Child Psychol Psychiatry 40, 649–659.

Howard, M.A., Cowell, P.E., Boucher, J., Broks, P., Mayes, A., Farrant, A., Roberts, N., 2000. Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism. NeuroReport 11, 2931–2935.

Ingersoll, B., 2008. The effect of context on imitation skills in children with autism. Research in Autism Spectrum Disorders 2, 332–340.

Ingersoll, B., Schreibman, L., Tran, Q.H., 2003. Effect of Sensory Feedback on Immediate Object Imitation in Children with Autism. J Autism Dev Disord 33, 673–683.

Jack, R.E., Garrod, O.G.B., Schyns, P.G., 2014. Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time. Current Biology 24, 187–192. doi:10.1016/j.cub.2013.11.064

Kätsyri, J., Sams, M., 2008. The effect of dynamics on identifying basic emotions from synthetic and natural faces. International Journal of Human-Computer Studies 66, 233–242.

Kleinhans, N.M., Johnson, L.C., Richards, T., 2009. Reduced neural habituation in the amygdala and social impairments in autism spectrum disorders. American Journal of ….

Lindner, J.L., Rosén, L.A., 2006. Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. J Autism Dev Disord 36, 769–777.

Loveland, K.A., Tunali-Kotoski, B., Chen, Y.R., Ortegon, J., Pearson, D.A., Brelsford, K.A., Gibbs, M.C., 1997. Emotion recognition in autism: verbal and nonverbal information. Dev Psychopathol 9, 579–593.

McDuffie, A., Turner, L., Stone, W., Yoder, P., Wolery, M., Ulman, T., 2007. Developmental correlates of different types of motor imitation in young children with autism spectrum disorders. J Autism Dev Disord 37, 401–412.

McIntosh, D.N., Decker, A.R., 2006. When the social mirror breaks: deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. Developmental Science 9, 295–302.

Ohta, D.M., 1987. Cognitive disorders of infantile autism: A study employing the WISC, spatial relationship conceptualization, and Gesture Imitations. J Autism Dev Disord 17, 45–62.

Pelphrey, K.A., Morris, J.P., McCarthy, G., LaBar, K.S., 2007. Perception of dynamic changes in facial affect and identity in autism. Social Cognitive and Affective Neuroscience 2, 140–149.

Piggot, J., Kwon, H., Mobbs, D., Blasey, C., Lotspeich, L., Menon, V., Bookheimer, S., Reiss, A.L., 2004. Emotional attribution in high-functioning individuals with autistic spectrum disorder: a functional imaging study. J Am Acad Child Adolesc Psychiatry 43, 473–480.

Reinl, M., Bartels, A., 2015. Perception of temporal asymmetries in dynamic facial expressions. Front. Psychology 6, 1107.

Reinl, M., Bartels, A., 2014. Face processing regions are sensitive to distinct aspects of temporal sequence in facial dynamics. NeuroImage 102 Pt 2, 407–415.

Rogers, S.J., Bennetto, L., McEvoy, R., Pennington, B.F., 1996. Imitation and Pantomime in High-Functioning Adolescents with Autism Spectrum Disorders. Child Dev 67, 2060.

Rogers, S.J., Hepburn, S.L., Stackhouse, T., Wehner, E., 2003. Imitation performance in toddlers with autism and those with other developmental disorders. Journal of Child Psychology and Psychiatry 44, 763–781.

Rump, K.M., Giovannelli, J.L., Minshew, N.J., Strauss, M.S., 2009. The development of emotion recognition in individuals with autism. Child Dev 80, 1434–1447.

Rutherford, M.D., Towns, A.M., 2008. Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. J Autism Dev Disord 38, 1371–1381.

Scambler, D.J., Hepburn, S., Rutherford, M.D., Wehner, E.A., Rogers, S.J., 2007. Emotional Responsivity in Children with Autism, Children with Other Developmental Disabilities, and Children with Typical Development. J Autism Dev Disord 37, 553–563.

Stone, W.L., Ousley, O.Y., Littleford, C.D., 1997. Motor Imitation in Young Children with Autism: What's the Object? J Abnorm Child Psychol 25, 475–485.

Tardif, C., Lainé, F., Rodriguez, M., Gepner, B., 2007. Slowing down presentation of facial movements and vocal sounds enhances facial expression recognition and induces facial-vocal imitation in children with autism. J Autism Dev Disord 37, 1469–1484.

Teunisse, J.P., Gelder, B.D., 1994. Do autistics have a generalized face processing deficit? International Journal of Neuroscience.

Walsh, J.A., Vida, M.D., Rutherford, M.D., 2014. Strategies for perceiving facial expressions in adults with autism spectrum disorder. J Autism Dev Disord 44, 1018–1026.

Wehrle, T., Kaiser, S., Schmidt, S., Scherer, K.R., 2000. Studying the dynamics of emotional expression using synthesized facial muscle movements. J Pers Soc Psychol 78, 105–119.

Weyers, P., Mühlberger, A., Hefele, C., Pauli, P., 2006. Electromyographic responses to static and dynamic avatar emotional facial expressions. Psychophysiology 43, 450–453.

Yirmiya, N., Kasari, C., Sigman, M., Mundy, P., 1989. Facial expressions of affect in autistic, mentally retarded and normal children. J Child Psychol Psychiatry 30, 725–735.