

Bangor University

DOCTOR OF PHILOSOPHY

Towards a model of human body perception

Gandolfo, Marco

Award date:
2020

Awarding institution:
Bangor University

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 13. Mar. 2024

Bangor University

DOCTOR OF PHILOSOPHY

Towards a model of human body perception

Gandolfo, Marco

Award date:
2020

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 12. Feb. 2020



PRIFYSGOL
BANGOR
UNIVERSITY

Towards a model of human body perception

Marco Gandolfo

Thesis submitted to the School of Psychology, Bangor University, in partial fulfilment
of the requirements for the degree of Doctor of Philosophy

Bangor University, Bangor, UK, December 2019

Yr wyf drwy hyn yn datgan mai canlyniad fy ymchwil fy hun yw'r thesis hwn, ac eithrio lle nodir yn wahanol. Caiff ffynonellau eraill eu cydnabod gan droednodiadau yn rhoi cyfeiriadau eglur. Nid yw sylwedd y gwaith hwn wedi cael ei dderbyn o'r blaen ar gyfer unrhyw radd, ac nid yw'n cael ei gyflwyno ar yr un pryd mewn ymgeisiaeth am unrhyw radd oni bai ei fod, fel y cytunwyd gan y Brifysgol, am gymwysterau deuol cymeradwy.

I hereby declare that this thesis is the results of my own investigations, except where otherwise stated. All other sources are acknowledged by bibliographic references. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless, as agreed by the University, for approved dual awards.

Towards a model of human body perception

Acknowledgements

This thesis is the result of three years in which I gained a lot of life experience. Before starting this path, I had a passion for cognitive neuroscience and experimental psychology. At present, this passion holds strongly but it is now scaffolded with a solid, unbiased, scientific approach. This is the main reason why I want to thank my first supervisor Paul Downing. His way to do research represents the best approximation to the scientific approach that I have ever seen: truly passionate, unbiased, and smart. Thanks for being the researcher that I will try to become in the future. I feel this might be possible thanks to your guidance. It was an honour to be supervised by such an excellent experimental psychologist.

I want to thank Aoife, for being my partner, for sharing adventures all over North Wales, for her presents, and for being close to me wholeheartedly. It is truly hard to accept that the passion for this job now separates us, hopefully not for long! I also want to thank you for your huge help and advice at work.

I want to thank my second supervisor Kami Koldewyn. She was an excellent example of good researcher and source of very useful feedback. I hope we will collaborate in the future. Furthermore, I also enjoyed her very good vegetables. I also want to thank Ken Valyear for being the chair of my committee and giving me useful advice.

I want to thank the beautiful friends I made here. Jon, it has been great to share with you so much food and booze thanks also for your precious advice. Ioana you are a nice storm. Can't forget Kohinoor, Thy, Ciaraness, Lucybus, Connybus, Maria Lucia, Emmabus, Simona, Julia, Katie. All of you have been my extended family here, and I am grateful to have spent some of my free time with you. I also want to thank my eternal roman friends: Mario, Adriano, Crosta, Santrino, Daniela, Super, Livia, Ferro, Lucio. I miss you a lot, but thank god you are always ready to party when I am back!

I want to thank the kitchen crew, particularly David Carey, Caroline Parkinson (she also fed me a lot), Rob, Giovanni, Gary, Becca, Patricia, Manon, Paul Mullins, Richard Binney. I had very nice discussions during the breaks, about science and about food (and its questionable quality!).

I want to thank the students I supervised: Lea, Lina, Sara, Ionela, and Poppy. You were of great help but also a great source of experience. I learnt a lot from you and I hope I was able to pass you some of my knowledge. And, of course, thanks to all of my patient participants!

I want to thank my former lab friends in Rome: Giusi, Ilaria, and Vanessa. It has been great to discuss with you about our research and to have your advice.

If we admit that is possible to thank inanimated entities, then I want to thank Pizza and Coffee. The former has been the funniest thing to cook, the latter my strongest link with my home and my savior in sleepy times.

Voglio ringraziare I miei genitori, per il loro supporto cieco verso le mie scelte, per il loro aiuto e il loro affetto, anche nei momenti difficili, anche essendo lontani. Spero un giorno di tornarvi ad essere geograficamente piu' vicino. Voglio ringraziare I miei fratelli, Paolo e Ale, siete fantastici e mi mancate molto. Voglio ringraziare mia nonna, e' sempre un piacere sentirti al telefono, anche per 2 minuti per un saluto urlante. Ringrazio anche nonna Concetta, mi manchi molto.

Summary

From just a glimpse of another person, we make inferences about their current states and longstanding traits. These inferences are normally spontaneous and effortless, yet they are crucial in shaping our impressions and behaviours towards other people.

What are the perceptual operations involved in the rapid extraction of socially relevant information? To answer this question, over the last decade the visual and cognitive neuroscience of social stimuli has received new inputs through emerging proposals of social vision approaches. Perhaps by function of these contributions, researchers have reached a certain degree of consensus over a standard model of face perception.

This thesis aims to extend social vision approaches to the case of human body perception. In doing so, it establishes the building blocks for a perceptual model of the human body which integrates the extraction of socially relevant information from the appearance of the body.

Using visual tasks, the data show that perceptual representations of the human body are sensitive to socially relevant information (e.g. sex, weight, emotional expression). Specifically, in the first empirical chapter I dissect the perceptual representations of body sex. Using a visual search paradigm, I demonstrate a differential and asymmetrical representation of sex from human body shape. In the second empirical chapter, using the Garner selective attention task, I show that the dimension of body sex is independent from the information of emotional body postures. Finally, in the third empirical chapter, I provide evidence that category selective visual brain regions, including the body selective region EBA, are directly involved in forming perceptual expectations towards incoming visual stimuli. Socially relevant information of the body might shape visual representations of the body by acting as a set of expectancies available to the observer during perceptual operations.

In the general discussion I address how the findings of the empirical chapters inform us about the perceptual encoding of human body shape. Further, I propose how these results provide the initial steps for a unified social vision model of human body perception. Finally, I advance the hypothesis that rapid social categorisation during perception is explained by mechanisms generally affecting the perceptual analysis of

Towards a model of human body perception

objects under naturalistic conditions (e.g. expectations-expertise) operating within the social domain.

Contents

Acknowledgements.....	I
Summary.....	III
Contents.....	V
Chapter 1: Introduction.....	1
1.1 History and background.....	1
1.1.1 Person perception.....	1
1.1.2 The functional question in vision science.....	5
1.1.3 Social vision.....	9
1.2 Social Vision: Faces.....	12
1.2.1 Face perception: Behavioural evidence.....	12
1.2.2 Face perception in the brain.....	17
1.3 Social vision: Bodies.....	22
1.3.1 Body perception: Behavioural evidence.....	24
1.3.2 Body perception in the brain.....	29
1.3.2.1 Viewpoint and Identity.....	32
1.3.2.2 Body in motion and in action.....	34
1.3.2.3 Body emotion.....	36
1.3.2.4 Extracting social categories from the body.....	37
1.3.2.5 Current models of human body perception.....	38
1.4 Overview of the thesis.....	43
Chapter 2: Visual perception of sex from human body shape is asymmetrical.....	45
2.1 Abstract.....	46
2.2 Introduction.....	47
2.3 Experiment 1.....	50
2.3.1 Introduction.....	50
2.3.2 Method.....	50
2.3.2.1 Participants.....	50
2.3.2.2 Apparatus and Stimuli.....	51
2.3.2.3 Design and Procedure.....	53
2.3.2.4 Data Analysis.....	53
2.3.3 Results.....	58
2.3.4 Discussion.....	58
2.4 Experiment 2.....	58
2.4.1 Introduction.....	58
2.4.2 Methods.....	59
2.4.3 Results.....	59
2.4.4 Discussion.....	60
2.5 Experiment 3.....	60
2.5.1 Introduction.....	60

2.5.2 Methods.....	61
2.5.3 Results.....	62
2.5.4 Discussion.....	62
2.6 Experiment 4.....	63
2.6.1 Introduction.....	63
2.6.2 Methods.....	63
2.6.3 Results.....	64
2.6.4 Discussion.....	64
2.7 Experiment 5.....	65
2.7.1 Introduction.....	65
2.7.2 Methods.....	66
2.7.3 Results.....	66
2.7.4 Discussion.....	66
2.8 Experiment 6.....	67
2.8.1 Introduction.....	67
2.8.2 Methods.....	68
2.8.3 Results.....	68
2.8.4 Discussion.....	68
2.9 General Discussion.....	69
Chapter 3: Perceiving emotion and sex from the body: evidence from the Garner task for independent processes.....	75
3.1 Abstract.....	76
3.2 Introduction.....	77
3.3 Methods.....	80
3.3.1 Participants.....	80
3.3.2 Stimuli.....	80
3.3.3 Design and Procedure.....	81
3.3.4 Analyses.....	82
3.4 Results.....	84
3.4.1 Experiment 1: Angry and Fearful.....	84
3.4.2 Experiment 2: Happy and Angry.....	85
3.4.3 Experiment 3: Happy and Sad.....	86
3.5 General Discussion.....	87
3.6 Concluding remarks.....	90
Chapter 4: Causal evidence for Expression of Perceptual Expectations in Category-selective Extrastriate Regions.....	92
4.1 Summary.....	93
4.2 Results and Discussion.....	94
4.3 Methods.....	100
4.3.1 Participants: Behaviour-only.....	100

Towards a model of human body perception

4.3.2 Participants: TMS.....	100
4.3.3 Imaging.....	101
4.3.4 TMS stimulation.....	102
4.3.5 Stimulus creation.....	102
4.3.6 Stimulus presentation.....	103
4.3.7 Task Procedures.....	103
4.4 Statistical analyses.....	104
4.4.1 Analysis.....	104
4.5 Data and Software Availability.....	104
Chapter 5: General Discussion.....	105
5.1 Summary of findings.....	105
5.2 Towards a model of human body perception.....	107
5.3 Encoding of sex from body shape.....	108
5.3.1 Encoding of body sex in the brain.....	113
5.3.2 Sex asymmetry across stimulus classes and future directions.....	115
5.4 Extracting multiple social cues from the body.....	117
5.5 Perceiving the body in a social world.....	125
5.6 Concluding remarks.....	128
References.....	131
Supplementary Materials.....	174

1. Introduction

1.1 History and background

1.1.1 Person perception

During the course of their life, humans develop the sophisticated ability to navigate into their social world. From the moment we are born, till the end of life, our social sphere becomes increasingly complex and, in parallel, our social knowledge about others expands considerably. Indeed, we collect, organize, and store information about other people to make use of it in our daily life when we encounter others. For example, when getting on a busy train, we can quickly infer the sex, race, age, emotional status of the people around us. Based on this categorical information that we extract very rapidly, we form expectations about them, about the job they might do, the place they are going, or the way they are feeling. Most importantly, such expectations can eventually shape our behavior towards them, for example by deciding to sit next to them or not, or to start a conversation. While expectations based on the quick extraction of social categories are not exempt to errors, leading to stereotypes, there is no doubt that this process is strategical and necessary to make sense or to react promptly to our complex social environment (Fiske and Taylor, 1984; Brewer et al., 1991).

Over the last half-century, social psychology has extensively studied these processes under the term of “Person Perception”. This field explored how we make use of well-formed expectations about others (Olson et al., 1996), how we summarize them into social categories (Macrae and Bodenhausen, 2000), and how these categorization processes guide inferences toward others (Fiske and Neuberg, 1990). However, forming impressions about others is not purely the result of inferences based on previous knowledge but is rather the result of an interaction between our well-established social knowledge and incoming sensory information from the surrounding environment. Yet, standard definitions of this influential field in social psychology define person perception as: “The processes by which people think about, appraise and evaluate other people” (APA Dictionary of Psychology, 2007). Surprisingly, even such basic definition lacks attention on one of the core elements of its own phrase: “Perception”. This contrast between the term and its definition did not go unnoticed. Gilbert (1998) started his chapter on the handbook of social psychology with a provocative statement (p 89):

“The phrase [Person Perception] appears to describe both an activity (perception) and the object toward that activity (person), but the object is described inadequately and the activity incorrectly.”

Such a strong stance evoked the proposal of renaming the discipline: “Ordinary Personology”. The aim was to define a more comprehensive term which highlighted the systematic investigation and understanding of attributes, such as temporary states (emotions, intentions), and enduring dispositions (beliefs, traits, and abilities) of other people.

But the phrase “Person Perception” hardly comes from nowhere. Social psychologists initially treated the understanding of persons as a matter of the perceiver’s *accuracy*. In agreement with Brunswick’s lens model (1956) external objects (distal stimuli, e.g. a person walking on the street) have properties that manifest as visual information (media, e.g. a smile) which may or may not be received by the observer (proximal stimulus) and if received they may or may not correctly interpreted (e.g. greeting). This “objective” approach (Gilbert, 1998) to the discipline, which emerged in parallel with Skinner’s radical behaviorism theories (Skinner, 1953), had to deal with the challenge of measuring and assessing the accuracy of context-independent personality traits. Indeed, in line with the success of intelligence testing in the cognitive domain, social psychologists rushed to the task of measuring individual differences in accuracy in person perception aiming to differentiate people in their ability to know the social world surrounding them. In this view, participant’s accuracy in judging others’ traits was compared with an operational criterion (what a person is really like). This approach raised at least two methodological difficulties. The first concerned the operationalization of the criterion: personality traits are not as simple to measure as other objects and, also, are subject to changes over time. Second, as famously noted by Cronbach (1955; 1958), accuracy scores were composed of multiple components and many of the individual differences were better explained by the decision bias of the participants across all the personality traits they rated (Kenny and Albright, 1987). These issues generated a large set of incoherent findings and led to the gradual loss of interest in such “empirical enthusiasm”, which lacked a strong theoretical basis (Bruner & Tagiuri, 1954) (but see Kenny, 1994, Trope, 1986; Ambady et al., 2000 for new brunswickian approaches).

From that time on, social psychology became largely interested in the *how* of person perception, particularly the processes by which judgments of people are made, and how impressions about others were formed. In fewer words, the focus of the field shifted to those cognitive processes that generated biased accuracy in trait judgments (Kenny and Albright, 1987). By the end of the 1970s, social psychology fully embraced the more rigorous methodologies brought by the cognitive revolution and, to some extent, embraced the assumption that any theoretical model of social perception should have proceeded along the more general models of human cognition (Gilbert, 1998), despite some concerns of independence as a discipline (Markus & Zajonc, 1985).

Social psychologists and person perception researchers in the following decades turned their attention on the errors of human judgments (Tversky & Kahneman, 1974; Nisbett and Ross, 1980). For example, Kelley (1950) showed how labelling a communicator as “cold” could cause students to dislike their lectures. In another study, participants were unconsciously exposed to positive or negative trait terms (e.g. “reckless”, “persistent”) before reading a description of a stimulus person. The mere activation of these categories through priming, affected further trait judgements based on verbal descriptions of other people (Higgins et al., 1977). These errors, when forming impressions about other people, are based on the comparison between short or long-term expectations built over previous knowledge together with judgments and inferences applied to new encounters (Bruner, 1957; Bruner and Postman, 1949). Although long-term expectations are generally adaptive, investigating when these undermine accuracy provides a useful framework to understand how impressions are formed (Asch, 1946; Higgins and Bargh, 1987).

Why do we make use of expectations when forming impressions about others? The social information we are exposed to, and that we need to make sense of in daily life, is indeed highly complex. To accurately respond to the environment, humans need to be equipped with stable internal representations of the environments in which they operate (e.g. schemas, mental models, expectations) (Johnson-Laird, 1983, McClelland, 1995). However, perhaps the most distinctive ability of human mind (Johnston & Hawley, 1994) is to respond flexibly to the presence of unexpected, novel information brought by others, and to adapt their behaviour accordingly. While these two skills are apparently mutually exclusive, unexpected and novel behaviours take place only if in the presence of stable long-term expectations. In this sense, people are able to

form impressions about others very rapidly because they make use of their prior expectations “schematized” in the form of social categories. Accordingly, the person perception processes gain their distinctive flexibility and efficiency in new situations only through making use of stable categorization processes being active in novel situations (Macrae et al., 1999; Macrae & Bodenhausen, 2000; Hastie & Kumar, 1979).

When encountering others, with the aim of simplifying the amount of information they are exposed to, people do not form an impression of another person as a unique, complex array of attributes, but rather they build their impressions by categorizing the person along social categories (sex, age, race etc.). This shapes the “perception” of others in at least two important ways. Firstly, perceivers may use the activated categorical knowledge to guide the encoding of any target-related information influencing the way we process others (Macrae, Stangor and Milne, 1994; Bodenhausen, 1988; Dovidio, 1986; Macrae and Bodenhausen, 2000). Secondly, perceivers may use the content of categorical knowledge to derive evaluations and impressions of a target, more commonly conceived in the form of stereotypical judgments (Fiske & Neuberg, 1990; Allport, 1954; Brewer, 1988). For example, when participants are primed with the (race) category “black”, they are faster in responding to category associated traits (e.g. “musical”), showing that once a certain category is activated participants gain rapid access to its contents (Dovidio, 1986). However, such rapid associations, even when primed implicitly, although they facilitate processing of compatible information, are largely dissociated from explicit evaluations as measured by self-reported racial prejudice (Devine, 1989; Dovidio et al., 1997). In turn, this explicit measure predicted relative evaluations of black and white interaction partners (Dovidio et al., 1997).

The work in the respective fields of social psychology and in person perception mainly focused on post-perceptual operations, particularly on how category-based knowledge (or expectations, stereotypes) would impact impressions, memories and actions towards other individuals (Brewer, 1991; Bodenhausen & Macrae, 1998; Fiske & Neuberg, 1990). Two factors particularly led the field to overlook the perceptual operations. Firstly, theoretically, perception had been conceived in a radically gestaltist sense, where pre-existing knowledge is fully integrated in a whole with incoming sensory data. This stance had the paradoxical effect which led researchers to be more concerned with the higher-level processes (such as judgments and beliefs) than with

perceptual processes (Gilbert, 1998). Secondly, methodologically, since the earliest days of the discipline (Asch, 1946), a favoured technique has been to present participants with verbal stimuli (e.g. category labels, trait adjectives). For example, a commonly used paradigm would consist of participants reading a list of traits, and forming an impression of the actor whom those traits verbally describe. While this was a successful, convenient, and well controlled method, it did not help to elucidate the perceptual operations by which these inferences are enabled (but see Zebrowitz MacArthur and Reuben Baron, 1983). As correctly pointed out by Macrae and Quadflieg (2010), previous work in “person perception” in social psychology has focused on the perception of personality (or on personology, to use Gilbert’s, 1998 expression) and not on the perception of people.

To explore this further, I will first review some basic foundations of perception, with a particular focus on vision science and its main overarching research questions. After, I will consider how vision science has been applied to the questions of ‘person perception’.

1.1.2 The functional question in vision science

Vision science has perhaps produced the most successful line of studies that demonstrated direct mapping between neuronal populations and visual features (Hubel & Wiesel, 1959,1962; Allman & Kaas, 1974; Van Essen and Maunsell, 1983; Zeki,1978). The disproportionate amount of findings in this direction provide an apparent support for a view of perception where stimuli cause their response directly, via projection mechanisms onto the nerves. However, insights from psychophysics show that the relationship between measurable physical stimuli and their correspondent human perceptions are not linear. For example, the light intensity of a visual stimulus can be varied, causing it to appear brighter, but the percept of this increase does not linearly change together with the luminosity of the stimuli (Fechner, 1860). Von Helmholtz, (1867) in his theory of perception posited that information available from the retina and other sensory organs is not sufficient to reconstruct the world; he famously noted that the mind makes a series of “unconscious inferences” to construct a coherent picture of its experiences. It is because of such inferences that, for example, observers perceive objects as having a constant size at various distances despite variation in the

visual angle (size constancy). Indeed, since the earliest days of its history, vision science has been interested in vision errors to support the existence of the “unconscious inferences”. Visual illusions, such as Kanisza Triangle (Kanisza, 1955), Ponzo (Ponzo, 1911), have been used to illustrate these errors, and show that estimates of visual features (color, brightness, length, size) can drastically change depending on context. Illusions shed a light on important properties of the visual system, such as perceptual grouping and the phenomenon of figure ground, and reveal that the visual system does something more than directly receiving light through the retina. Visual illusions occur in special circumstances and, while revealing the “nuts and bolts” of the visual system, remained far from providing a full understanding how everyday perception might work and what its final goal is (Nakayama, 1994).

The importance of understanding how vision operates in its context was not unnoticed. Gibson (1966) in his theory of ecological optics, had the intuition that many of the visual regularities in the optic array convey useful information about the world. His theory assumed that perception serves an adaptive function and thus guides biologically and socially functional behaviours (Zebrowitz MacArthur & Baron, 1983) as, for instance, supporting action execution (Gibson, 1979). Inspired by Gibson, Marr (1980; 1982) proposed vision as a process constituted by a set of sequential stages. In his view, before understanding the formal properties of the computations (algorithmic level) and its physiological substrates (implementation level), it is crucial to understand the goal of the computation (computational level) that the visual system performs.

In physiology, vision scientists originally believed that visual processing in the brain was confined to the striate cortex. However, starting from the 1970s neurophysiologists identified numerous additional visual regions in the posterior part of the brain (Allman & Kaas, 1974; Van Essen, 1985; Zeki, 1978; Gross, Rocha-Miranda and Bender, 1972). Years later, neuropsychological and functional magnetic resonance imaging (fMRI) studies have extensively demonstrated that human and primate ventral temporal cortex show striking dissociations beyond simple visual features: starting from classic large-scale dissociations between living and non-living objects found on neuropsychological patients (Warrington and Shallice, 1984), until a more fine-grained categorical organization with regions responding selectively to faces, bodies, tools, scenes, words, biological motion demonstrated with fMRI in healthy participants (Kanwisher et al., 1997; Sacchett and Humphreys, 1992; Barton et al., 2002; Downing,

2001; Peelen & Downing, 2005; 2007; Epstein & Kanwisher, 1998; Epstein, 2008; McCandliss et al., 2003; Grossman et al., 2000). This huge allocation of brain anatomy towards vision suggests that its role is more important than reading out simple visual features but rather serves myriad of functions.

Most vision researchers now agree that visual perception is more than just receiving an image on the retina, and that it is clearly related to understanding an image. However, there is less agreement on what is to be understood (Adams et al., 2011). Inspired by Marr (1982) and supported by studies with lesion patients (Goodale & Milner, 1992), it is suggested that the primary goal of ventral visual regions is to obtain view-invariant object descriptors that permit object identification (the highest stage for Marr). Following this view, it is possible that categorical preferences, including the one for the social stimuli (Caldara, 2006), may be reduced to preferences for the visual properties that are characteristic of those categories (a convergence of low level features), even in the absence of overt category recognition (Rice et al., 2013; Andrews et al., 2015; Op de Beeck et al., 2008; Ishai et al., 1999).

Is stimulus recognition the only task that the visual system performs? Studies have demonstrated that high level visual areas encode representational, functional information about their preferred category even when the shape and low-level features of their preferred stimuli are matched (Proklova et al., 2016; Kaiser et al., 2016; Macdonald & Culham, 2015). In parallel, a number of studies have shown that the right fusiform gyrus- a region that respond to social stimuli as faces and bodies – is activated by simple geometric shapes (highly different from their preferred stimuli) when these move in a way akin to social agents (Gobbini et al., 2007; Martin & Weisberg, 2003; Castelli et al., 2000). These findings show that occipito-temporal visual regions do not contain only view-invariant representations of specific categories of stimuli but rather convey something about the knowledge of these categories for the observer (Bracci & Op de Beeck, 2016; but see Bracci et al., 2019).

This view is further corroborated by a set of intriguing findings which challenge the idea that occipito-temporal category selective regions are purely visual. Functional MRI studies show that category selectivity is still present, although weaker (Op de Beeck et al., 2019), in congenitally blind individuals for objects presented in non-visual modalities (e.g. tactile or auditory) (Bi et al., 2016, Van den Hurk, 2017; Peelen et al.,

2014; Kitada et al., 2014). For example, haptic body perception induced activity in the body-selective region EBA in blind participants (Kitada et al., 2014). Whilst more research is needed to understand the nature of these representations, these findings indicate that vision might only be the most reliable source from which these representations are formed, but that these can also be acquired in absence of visual experience.

Taken together these findings suggest that regions of the ventral temporal cortex may encode properties of their preferred categories which goes beyond view-invariant object recognition (Peelen and Downing, 2017). Before the advent of brain imaging, Gibson (1966) had the intuition that vision could only be understood investigating the specific role that it plays within broader domain-specific networks (such as the support for actions, Gibson 1979). More recently, Peelen and Downing (2017) proposed that the goal of visual regions in the ventral temporal cortex is to form cross-modal perceptual representations that contribute to the efficient performance of a range of tasks that humans perform in daily life such as social cognition, reading, navigation, and tool use. In this sense, these regions encode fundamental task-relevant representations independent from each other and that serve functional and domain specific brain networks such as person recognition, tool use, understanding others' actions and emotions, and navigation.

The current and future challenges for vision scientists are likely entangled in this question: what are the functions that vision supports beyond object recognition? (Peelen & Downing, 2017). It is only by considering vision in the context of the functions that it serves that we can better understand its fundamental mechanisms. There are two possible ways in which this can be achieved: by using more realistic, ecologically valid paradigms (Hasson & Frith, 2016; Quadflieg and Koldewyn, 2017; Peelen and Kastner, 2014), and/or by integrating concepts and methods used in vision research together with the insights from other disciplines (Adams et al., 2011). For example, when considering the domain of social cognition, how are visual regions that process social stimuli integrated with the knowledge that we have about others? It is only with a broader view of the cognitive operations that other social brain regions perform that it will be possible to understand the functional contribution of visual areas responding to social stimuli.

1.1.3 Social Vision

Humphrey, a primate vision scientist, in 1976 stated: “Experimental psychologists in Britain have tended to regard social psychology as a poor country cousin of their subject[.]”. At a time when the cognitive revolution was taking place and social psychology was borrowing paradigms and concepts from it, Humphrey (1976) provocatively advanced the hypothesis that the brain evolved as a consequence of the demands of the social environment. In his view, the creative intellect, typical only of the most evolved species, developed from the need to deal with social interactions. Being ahead of his times, Humphrey was not wrong in anticipating that in the next decades the social stimuli, and particularly faces, would have become of great interest for vision scientists investigating both humans (Bruce and Young, 1986; Haxby et al., 2000; Haxby et al., 2002; Sergent et al., 1992; Barton et al., 2002) and non-human primates (Tsao et al., 2003; Tsao et al., 2006; Perret et al., 1984; Perret et al., 1992).

The respective fields of social psychology and vision science remained separate for many years despite their converging interests. On one hand, vision science has used social stimuli for decades, for example faces (Bruce and Young, 1986; Duchaine and Yovel, 2015), bodies (Peelen and Downing, 2007), and biological motion (Johansson, 1973; Bonda et al., 1996; Grossman et al., 2000). However, the field is now challenged by questions that go beyond the analysis of shape features of the stimuli or object recognition (Marr, 1980; Peelen & Downing, 2017). Modern research in visual perception supports the view that conscious perception does not consist in a copy of a retinal image, but it is the result of constructive processes (Bruner, 1957). Here, multiple factors such as personality, motivation, cognitive set and expectations play a role in the perception of the social stimuli. On the other hand, it would appear that the field of social psychology has come to appreciate how effortlessly social categorisation takes place (Bargh et al., 1999; Macrae & Quadflieg, 2010; Freeman & Johnson, 2016) and how this is entangled in processing of simple visual cues (Macrae & Martin, 2007; Martin & Macrae, 2007). In doing so, modern social psychological research has been drawing the necessary attention to the importance of perceptual operations in how we form impressions about other people.

The mutual interests of social psychology and vision science were integrated with the recent proposal of a “Social Vision” approach (Adams et al., 2011). Social vision

aims to highly integrate the insights and methodologies from both disciplines. Vision science has a lot to offer in terms of techniques and concepts that have been used to study basic social stimulus processing. For example, with high experimental control it is possible to manipulate the socially relevant dimensions embedded in the perceptual features of stimuli (e.g. gender, race, identity) on a continuous scale. With such manipulation, for example, it is possible to investigate through visual adaptation, whether specific neural populations within the visual system are tuned to these features. Similarly, social psychology provides theoretical distinctions which have proven to be useful for memory retrieval, such as categorization and individuation (Macrae and Bodenhausen, 2000). These concepts can offer vision science the opportunity to expand known perceptual models (e.g. the dual-process model for faces – Bruce and Young, 1986). For example, social categories triggered by distinct visual cues (Race – color; Sex – shape) can offer new inputs to perceptual models. Further, social cues convey visual information that cuts across the classic static versus dynamic distinction in dual route models (Haxby et al., 2000). For example, we can efficiently extract sex information from neurally dissociable body motion and body shape (Johnson and Tassinary, 2005). Importantly, social vision approaches assume that social perception is based on domain-specific processes and functionally specialised mechanisms of the cognitive system (Atkinson et al., 2011 from Adams et al., 2011). In sum, an interdisciplinary social vision discipline can be used to address the fundamental functional question “what is the visual system for?” in the domain of Social Cognition (Peelen and Downing, 2017).

In line with a social vision approach, a “Dynamic Interactive model” (Freeman and Ambady, 2011; Freeman and Johnson, 2016) has been proposed. This computational model aims to explain social categorisation at a perceptual level by positing a continuous interaction between bottom-up sensory information and top-down information (e.g. task demand, goals and stereotypes). The model includes four main levels: 1) A cue level, containing a set of detectors of visual features (facial and bodily cues, e.g. large jaw) and auditory features (vocal cues, e.g. low pitch), which are directly stimulated by bottom-up sensory information from other people; 2) A category level, that contains different static and dynamic category pools (e.g. sex, age, race, emotional expression); 3) A stereotype level, that includes all category-related stereotypes; 4) A higher order level, that includes Task Demands (e.g. sex categorization), processing

goals, or motivations. In such a model, perceptions of other people gradually emerge through ongoing cycles of interactions between the four levels and within different nodes of each level. For example, if a male face is presented, direct stimulation of the male cues at the cue level will facilitate the “Male” category in the category level and at the same time inhibit the category “Female”. In turn, the category representation will excite the category-related stereotype “Aggressive” together with the category “Black”, which in turn will further facilitate the category “Male” and so on and so forth until stable person construal is achieved. This model provides a unified example of how the two literatures can be combined, by describing how the lower-level perceptual processing modelled in the cognitive literature (Bruce & Young, 1986; Valentin et al., 1994) works in concert with the higher-order social cognitive processes (about how knowledge about individuals and groups is learned, stored and accessed, Bodenhausen and Macrae, 1998; Fiske and Neuberg, 1990).

The assumption that initial percepts are sensitive to social dimensions/categories and to top-down factors (e.g. stereotypes, expectations, attitudes) independently from top-down feedback provides possible new lines of investigations for vision science. Recent work based on this assumption has provided evidence that typical behavioural paradigms that revealed tuning of neural populations to specific visual features (e.g. visual adaptation) can also show effects for socially relevant information conveyed by the face (Leopold et al., 2001; Webster and Macleod, 2011), the body (Winkler and Rhodes, 2005; Brooks et al., 2019a; Brooks et al., 2018) and biological motion (Troje et al., 2006). In parallel, a series of fMRI studies showed evidence that the fusiform gyrus, responsive to face and face-like stimuli, is sensitive to categorical distinctions of identity (Rothstein et al., 2005), gender (Freeman et al., 2009; Contreras et al., 2013), race (Brosch et al., 2013; Contreras et al., 2013; Kaul et al., 2012), emotion categories (Wegrzyn et al., 2015), and of the co-activation of simultaneous multiple social categories (e.g. sex and race; Stoler and Freeman, 2017; Stoler and Freeman, 2016; Bagnis et al., 2019). Further, evidence suggests that body- and face- selective cortices (EBA and FFA, respectively) can also be shaped by higher order social cognitive processes such as ingroup-outgroup effects (Vanbavel et al., 2008) and gender stereotypes (Quadflieg et al., 2011).

In which way can a vision scientist interested in social stimuli make use of social vision approaches? Investigating the visual information conveyed by social categories

will likely contribute to the expansion of traditional neurocognitive models of face perception (Haxby et al., 2000; Bruce and Young, 1986; Duchaine and Yovel, 2015). Further, social categories may be important dimensions to better understand the neural and cognitive representations of human bodies, where, to date, a “standard” neurocognitive model has not yet been established. Moreover, adopting such an approach can draw interesting parallels across social stimulus type. It is possible that information conveyed by social cues (such as sex, age, race), equally present across bodies, faces, voices and biological motion, will show similar “cross-modal” neural and cognitive architecture. Does the dimension of sex share similar representations or cognitive structures across faces, bodies and voices? Such questions can help to inform and draw more comprehensive, ecologically valid models of person perception (Yovel and Belin, 2013; Freeman and Ambady, 2011; Yovel and O’toole, 2016).

1.2 Social Vision: Faces

Even a brief glance of a face is sufficient to provide information about sex, age, race, emotional status and identity of others. Because of this extraordinary ability that humans have in extracting such information, the human face has generated a huge interest among psychologists. As such, the human face has been so extensively researched that recent person perception models under the framework of social vision have been conceived considering solely (or mostly) findings with face stimuli (Freeman and Johnson, 2016; Adams and Kveraga, 2015). Thus, the objective of the following section is twofold: first, to provide an overview of the main behavioural and brain findings concerning face perception, with a particular focus on those paradigms that have proven useful in defining current perceptual models of face perception; second, to illustrate how social vision approaches provide a highly useful contribution for understanding the visual processing of the face.

1.2.1 Face perception: Behavioural evidence

Faces attract most of our attention when we perceive other people within a naturalistic scene (Fletcher-Watson et al., 2008). Such preference for face-like stimuli appears early, in the first 3 months of life (Morton and Johnson, 1991; Simion, Leo, Turati, Valenza, Dalla Barba, 2007; Mondloch et al., 1999). The special attention-

capturing quality of faces has been shown using several paradigms. Faces capture attention in visual search among other objects (Herschler and Hochstein, 2005) even when they are task irrelevant (Langton et al., 2008). Changes for faces are detected more rapidly and accurately than changes for other common objects (Ro, Russel, and Lavie, 2001), and schematic faces, particularly when depicted in a typical “smile” icon, are more resistant to inattention blindness (Mack and Rock, 1998).

The disproportionate amount of attention deployed to faces is reduced, or even abolished, when faces are not arranged in their typical configuration. That is, when faces are not presented in their prototypical spatial arrangement (two lateralised eyes over a centralised nose over a mouth) they lose their “privileged” status during stimulus processing as evidenced by *face inversion effects* (Yin, 1969; Maurer et al., 2002). When faces and objects are inverted, recognition performance is impeded more dramatically for faces than for objects. Such an effect indicates that perceivers not only extract information relative to single facial features (e.g. face parts such as mouth, nose and eyes) but also encode information related to the arrangement of such features. Further, when distorting the relations among the features of a face – such as the rotation of the eyes and mouth in the famous Thatcher illusion (Thompson, 1980) – a face looks grotesque only if it is presented upright. When rotated from upright adults see the face as increasingly less bizarre (Murray et al., 2000), possibly because people more accurately detect facial features when presented in a context of a whole upright face rather than in isolation or in the context of an inverted face (Tanaka and Farah, 1993). Finally, the holistic, Gestalt-like, processing of faces is also visible in the “*composite face effect*”. Subjects are slower and less accurate in recognizing the top-half of one face presented in a composite together with a bottom-half of a face of another person. Performance improves drastically when the top and bottom part are misaligned or the composite face is inverted (Young et al., 1987; Hole, 1994).

How is the information provided by faces encoded? Recognition memory for unfamiliar faces, rated similar to a prototype, is inferior compared to faces rated as unusual or distinctive in appearance (Light et al., 1979; Valentine and Bruce, 1986). However, if tasked with distinguishing typical from jumbled faces, distinctive faces take longer to be categorised as faces (Valentine and Bruce, 1986). These effects of distinctiveness on the recognition of faces are explained in terms of faces being encoded by reference to a facial prototype (Valentine, 1991; Rhodes et al., 1987).

These findings are in line with the proposal of a multidimensional similarity space in which face representations are organised (Valentine, 1991; Wenger and Townsend, 2001; Leopold et al., 2001). In a face space, each face is represented by a location in a psychological similarity space. Faces represented close are similar to each other, faces that are distant are more dissimilar. The dimensions of the space represent the attributes on which faces may vary. These attributes possibly correspond to mutually exclusive global properties of the face.

The crucial dimensions of the face space have been unravelled using visual adaptation paradigms (Webster and Macleod, 2011; Webster and MacLin, 1999). After constant stimulation of a particular stimulus characteristic (adaptation) the perceptual system shifts its norm, away from the mean, towards the adapted characteristic. Interestingly, this approach has been applied successfully not only to low level characteristics of the faces – such as the central distortion of features of a face (Webster and MacLin, 1999) or their size (Zhao and Chubb, 2001) - but also on high-level and ecologically relevant dimensions of the facial stimuli such as sex, ethnicity, emotional expression and identity (Webster et al., 2004; Leopold et al., 2001; Hsu and Young, 2004). For example, norm-based coding in a face-space was demonstrated by morphing together 200 faces together to produce a prototype face (Leopold et al., 2001). This was assumed to be the centre of the face-space. Each unique identity could be measured in terms of Euclidean distance from the prototype face. Adaptation to an anti-face (on the opposite end of the space from the face identity), led to misperception of face identity that was biased towards the corresponding face from which the anti-face adapter was derived. Similarly, such effects have also been found for more general, natural variations that define facial features typical of certain social categories. For instance, after adapting to a male face, an androgynous face (a morph between male and female face) appears to be more feminine. These findings show how adaptation paradigms are useful to reveal the architecture of the cognitive representations of the face and, in parallel, inform neural models of face processing by capturing the way different neural populations are tuned to face properties (Winston et al., 2004; Leopold et al., 2006).

Another open question concerns the relationship between different sources of information conveyed by the face. If faces are stimuli that vary along multiple dimensions, then it is reasonable to ask whether such dimensions interact with each

other (integrally processed), or they are instead processed independently from each other. The Garner selective attention task (Garner, 1974; Algom and Fitousi, 2016; Pomerantz and Garner, 1973; Garner, 1976) has been extensively used in the face literature to answer this question. The Garner selective attention task (Garner, 1974) is a speeded classification task where participants are asked to judge one dimension of a face (e.g. sex) while another task-irrelevant dimension is fixed (e.g. identity – *control block*), or is randomly varied on a trial-by-trial basis (*orthogonal block*). Lower performance, as indicated by longer reaction times, in the orthogonal compared to the control block indicates integral processing of the dimensions tested; participants cannot attend to one dimension while fully ignoring the other irrelevant dimension. Conversely, equal performance in the control and orthogonal conditions indicates separability of processing of the dimensions; it is possible to attend to one dimension while fully ignoring the irrelevant variation of the other.

The Garner selective attention task has been applied to test several predictions made by classic dual-route neural models of face perception (Haxby et al., 2000; Bruce and Young, 1986). For example, Ganel and Goshen-Gottstein (2002) found that irrelevant variation of sex interfered with identity judgments and irrelevant variation of identity interfered with sex judgments. Interestingly, this finding has challenged the idea of a *parallel-route hypothesis*, where identity information (face identification process) proceeds along a parallel pathway that is separable from face classification processes (which include sex processing) (Bruce and Young, 1986). Instead, the data are more consistent with fMRI models that do not directly posit neural separability of the processing of different type of static/invariant information from the face, such as sex and identity information (Haxby et al., 2000).

Other studies have tested the relationship between variable (emotional expression, gaze or facial speech) and invariant (sex or identity) characteristics of the face, in line with the proposal of perceptual independence between these two types of information (Schweineberger and Soukup, 1998; 1999; Atkinson et al., 2005; Le Gal and Bruce, 2002; Karnadewi and Lipp, 2011; Becker, 2017). Despite that the absence of interference (independent processing) was expected, they have consistently reported evidence for partially-integral processing of variable and invariant face information by showing asymmetric Garner interference patterns (Atkinson et al., 2005; Schweineberger and Soukup, 1998; 1999; Karnadewi and Lipp, 2011; Becker, 2017;

but see Le Gal and Bruce, 2002). Specifically, asymmetric interference was found in terms of irrelevant variation of invariant information (sex; identity) impacting judgments of variable information (speech; emotional expression) but not vice-versa.

Asymmetrical interferences are generally explained with reference to a parallel-contingent model of processing; for example, where judgments of emotional expression are impeded, resulting in an increased RT, by irrelevant variation of identity, but not vice-versa (Schweineberger and Soukoup, 1998;1999). The parallel-contingent model asserts that identity information shapes the ongoing processing of emotional expression, and thus is used as a reference to process the emotional expression. Other studies have shown that two variable characteristics of the face (eye gaze and emotional expression) are integrally processed, with a stronger influence of gaze processing on emotional expression judgments than the converse (Ganel et al., 2005; Graham and Labar, 2007; but see Ganel, 2011). Taken together, these findings have confirmed and/or challenged predictions made by existing cognitive models of face perception. Moreover, they demonstrate that the Garner paradigm is a useful tool that offers new insights on the predictions made by such models. For example, the Haxby et al. (2000) model does not directly specify the degree of functional separation between variant and invariant routes. However, Haxby et al. (2000) hypothesise the possibility of a one way interaction between the two routes where the invariant route plays a “supporting role” in the perception of variable information. Consistent evidence of asymmetrical Garner interference between variable and invariant information confirm the possibility that the cross-talk between these two routes is unidirectional or biased in one direction, from ventral-invariant to supratemporal-variable but not (or less) vice-versa (Atkinson et al., 2005).

The relationship among different sources of information conveyed by faces has also attracted the interest of social psychologists. A first distinction, recently applied to perceptual operations but classically studied in the context of memory retrieval, regards categorization and individuation processes (Macrae and Bodenhausen, 2000). Categorisation is suggested to impact semantic memory and account for classifying faces to a certain social category (e.g. sex, age, race), while individuation impacts episodic memory and accounts for unequivocal determination of a face as a unique ensemble of features and characteristics (identity). Cloutier and colleagues (2005) outline that these processes might rely on distinct perceptual operations. Across three

experiments they show that identity judgments (which require individuation) were more affected by classic manipulations of processing difficulty (i.e. inversion, blurring, and speeded presentation) compared to categorical judgments of sex. Following this distinction, other studies have demonstrated that category-based judgments of faces (i.e. sex judgments) primarily rely on the extraction of featural information from several, isolated, facial cues (Macrae and Martin, 2007; Brebner et al., 2009; Martin and Macrae, 2007; Yamaguchi et al., 2013; but see DeGutis et al., 2012; Baudouin et al., 2006; Bruce et al., 1993). Among these, one of the most reliable is hairstyle (Brown and Perrett, 1993; Burton et al., 1993; Nestor and Tarr, 2008; Goshen-Gottstein and Ganel, 2000). Priming a male or female hairstyle generates comparable congruency effects to a whole face when judging male and female names or masculine and feminine words (Macrae and Martin, 2005; 2007). Featural processing of simple cues is then a perceptual mechanism that can instantiate person categorisation (Macrae and Martin, 2007). The “simplicity” and speed of such perceptual operations might account for the preferred use of categorisation processes in higher order social categorisation processes that impact impression formation and stereotyping (Brewer, 1988; Fiske and Neuberg, 1990). Conversely, processing of configural information is required for individuation processes (i.e. identity judgments). When the configuration is disrupted via increasing the degree of rotation of a face from upright to inverted, identification performance gradually decreases (Cloutier and Macrae, 2007).

Taken together these findings show that seeing a face captures the attention of the observers in a unique way. Further, faces are a special perceptual object for the observers because they are perceived holistically, and are not only constituted by a sum of their parts. Faces also convey a rich amount of information that is possible to be summed up in several, different dimensions. The relation among these dimensions is of great interest to confirm or expand proposals coming from existing neurocognitive models of face perception. The social categorisation processes that determine the way in which we form an impression of others are constrained by the way such perceptual operations take place.

1.2.2 Face perception in the brain

The configural visual processing that takes place when we look at faces, and our outstanding ability to categorise them with a brief glance along with their social characteristics (e.g. identity, sex, age, emotional expression), rely on computations carried out in functionally specific face-selective brain regions. That is, face selective areas provide a likely neural locus for those visual processes that distinguish the way we perceive faces compared to other objects (e.g. inversion, composite effects etc.).

In the nonhuman primate brain, face selective patches have been originally reported in the inferotemporal cortex (IT) (Gross et al., 1972; DeSimone et al., 1984), in the superior temporal sulcus (STS) (Perrett et al., 1982; Rolls, 1984), and amygdala (Rolls, 1984; Rolls et al., 1985). Recently, studies combining fMRI and single cell recordings revealed that 97% of the neurons located in an fMRI-defined face-selective patch in the macaque brain are selective for faces. Similar percentages have been also found in other face selective patches localised with fMRI in the macaque brain (Freiwald and Tsao, 2010). Further, parametrised images of several different faces have been linearly reconstructed using responses of ~200 face-selective cells in three face patches (AM, anterior medial; MF, middle fundus; ML, middle lateral) of the macaque temporal lobe (LeChang and Tsao, 2017). The cell firing-rate in these patches was proportional to the projection of a certain face stimulus onto a single axis of a high dimensional linear face space obtained via principal component analysis (PCA) on the shape and appearance descriptors. This finding advances the idea that face identity is encoded by a linear metric coordinate system, in support of the validity of computational models that conceive face representations encoded into multidimensional linear spaces both in nonhuman primates and humans (Edwards et al., 1998; Freiwald et al., 2009; Valentine, 1991; Leopold et al., 2001).

In the human brain, the earliest studies which have proposed that face processing is dissociable from processing other objects came from observations of brain-damaged patients. Patients show specific impairment for faces compared to other objects (Hoff and Potzl, 1937; Ellis and Florence, 1990; Sergent and Signoret, 1992) or specific impairment for objects with face processing preserved (Moscovitch et al., 1997). Acquired lesions over different cerebral structures in the right hemisphere are linked with perceptual impairments for faces at a different level of severity: from deficits that

concern all the perceptual operations on faces but not on other objects (e.g. matching to identical views of the same face) to more fine-grained inability to derive configurational properties of the face (Sergent and Signoret, 1992).

Based on evidence coming from brain-damaged patients and cognitive studies, the first comprehensive cognitive model of face processing was proposed by Bruce and Young (1986). According to Bruce and Young (1986) processing of a face begins with the generation of a view-centered representation of the face. This representation, regardless of the familiarity of the face, is analysed by two independent modules, one for facial expressions and one for speech analysis. Moreover, such a structural representation of the face is compared with face-recognition units (FRUs). If the represented face matches with stored FRUs, then the face will be recognised. Familiar face recognition is based on a structural code that allows for recognition of faces regardless of changes in pose, expression, and illumination. Recognition of unfamiliar faces instead uses more simple directed visual processing (pictorial codes) that are more sensitive to changes in the view of the faces. Importantly, this model predicts that identity and expression information are processed independently (but see latest revisions of the model Calder and Young, 2005; Calder et al., 2001; Bruce and Young, 2011) and that the face classification processes for unfamiliar faces (e.g. sex classification) are also separable from face identification processes of familiar faces. The framework of Bruce and Young (1986) was the first to bring together the literature coming from several sources (patients, primates, and cognitive studies) and, even if not directly committed in investigating the issues of implementation, has had a continuing influence on face perception research (Schweineberger and Burton, 2011); for example, by proposing that face perception is achieved via independent systems working in concert.

The introduction of functional magnetic resonance imaging (fMRI) in the 1990s allowed for the investigation of the implementation of face perceptual processes in the healthy brain. Functional localiser approaches revealed a cluster of face-selective voxels in the fusiform gyrus (Fusiform face area, FFA) (Kanwisher et al., 1997). This individually defined region showed higher responses for faces than scrambled faces, houses or hands. Not long after, two other face-selective regions were found: the inferior occipital gyrus (occipital face area, OFA, Clark et al., 1997; Gauthier et al., 2000) and the posterior superior temporal sulcus (pSTS-FA, Hoffman and Haxby,

2000). More recently, additional face selective areas have been revealed in the anterior temporal lobe (ATL-FA, Anzelotti and Caramazza, 2017; Harry et al., 2016; Rajimehr et al., 2009), the anterior portion of the superior temporal sulcus (aSTS-FA, Pitcher et al., 2011), and the inferior frontal gyrus (IFG, Fox et al., 2009; Chan and Downing, 2011).

What aspects of face processing does each of these regions support? Using an orthogonal one back task, asking participants to detect the repetition of face identity or emotional expression, Haxby and Hoffman (2000) identified a “core system” for the visual analysis of the face. This is composed of the OFA, the FFA and the pSTS-FA (Haxby et al., 2000). In their influential neurocognitive model, Haxby et al. (2000) proposed that OFA, engaged at early stages of face processing, sends inputs to the FFA, where invariant aspects of faces (e.g. identity, gender) are represented. The OFA also sends input to the pSTS-FA, which instead represents variable aspects of the faces such as emotional expression, eye-gaze and speech. Their model also suggests the existence of an “extended system” of face perception, not directly visual but that supports operations of areas in the core system. The extended system is comprised of the amygdala and limbic system, the auditory cortex and the intraparietal sulcus; which receive inputs and send feedback to the pSTS-FA to support processing of emotion, speech and gaze direction, respectively. Moreover, the extended system also includes the anterior temporal lobes which receive inputs and send feedback to the FFA in support of decoding semantic information of familiar faces and processing of facial identity. The network of face selective regions operates over two streams: a ventral stream which mainly extracts form information from faces, and a dorsal stream, which is highly involved in processing dynamic information from faces.

Almost twenty years of research following the Haxby et al. (2000) model has not disconfirmed their framework, but rather has expanded or refined their proposal of two pathways for face processing (Duchaine and Yovel, 2015). Studies on patients have revised the idea that OFA is the only entry point of visual information for face-processing (Rossion et al., 2003; Sorger et al., 2007; Steeves et al., 2006). For example, patient DF, with damage to bilateral OFA showed unaltered activity of FFA and pSTS-FA (Steeves et al., 2006). Other fMRI studies, in line with the behavioural findings using the Garner selective attention task (Schweineberger and Soukoup, 1998; 1999; Atkinson et al., 2005), have suggested that FFA may also be involved in facial expression processing (Ganel et al., 2005; Xu and Biederman, 2010; Fox et al., 2009),

possibly because it is broadly sensitive to shape information (Said et al., 2011). Further, the use of dynamic face localisers revealed that OFA, FFA, and the pSTS-FA show higher response with dynamic stimuli (Fox, 2009; Pitcher et al., 2010). The pSTS-FA showed particularly high response in dynamic vs static stimuli and was found active together with two other dorsal regions engaged only with dynamic stimuli, the aSTS-FA (Pitcher et al., 2011) and the IFG (Fox et al., 2009; Pitcher et al., 2011; Chan and Downing, 2011). Finally, along the ventral pathway, brain-lesioned patients and fMRI studies have suggested the existence of a critical area for face processing in the ATL-FA (Busigny et al., 2014; Rajimehr et al., 2009; Anzelotti et al., 2013; Yang et al., 2016). These studies proposed that ATL-FA is engaged in identity processing. For example, only the ATL-FA compared to FFA and OFA showed repetition suppression to pairs of different images of the same celebrity. This adaptation effect was found also in a patient with acquired prosopagnosia patient in the absence of functioning right OFA and FFA (Yang et al., 2016; but see Axelrod and Yovel, 2015).

Social vision approaches are revealing how regions of the core and extended network outlined in the Haxby et al. (2000) model respond to social categorisation of faces (e.g. sex, age, race) and to the identification of personality traits (Kawakami et al., 2017; Freeman and Johnson, 2016; Quadflieg and Macrae, 2011; Bagnis et al., 2019). Categorical perception of sex and race of faces has been associated with activity in the FFA (Cloutier, Turk and Macrae, 2008; Contreras et al., 2011; Ng et al., 2006; Golby et al., 2001; Freeman et al., 2009). One challenge for such studies concern the extent to which such differential activation of fusiform regions in response to different social categories is not explained by mere systematic changes of shape and/or colour naturally linked with these categories. For example, subjective ratings of masculinity or femininity did not correlate with activation of the fusiform gyrus (while objective parameters of face gender did) but rather with activation of the orbito-frontal cortex (Freeman et al., 2009), a region that has been linked with sex and attractiveness judgments (Winston et al., 2007). Concerning personality traits, the amygdala has been linked both with the simple perception and explicit judgments of untrustworthiness (Adolphs et al., 1998; Winston et al., 2002), possibly because some personality judgments might rely on emotional processing. Furthermore, amygdala activation has also been linked to racial prejudice (Cunningham et al., 2004; Vanbavel et al., 2008; Phelps et al., 2000; Wheeler and Fiske, 2005; Kawakami et al., 2017) although the

extent of its involvement is likely modulated by task goals (Cunningham, Van Bavel, Johnsen, 2008). In sum, to understand how the regions within both the core and extended network of face processing contribute to social categorisation and trait attribution, further empirical attention is needed. One interesting direction is that, as predicted by the dynamic interactive model (Freeman and Ambady, 2011), even regions of the core system, involved in the visual analysis of the face, like the FFA, might be sensitive to social categorisation and stereotyping processes (Freeman and Johnson, 2016; Stoler and Freeman, 2016; Stoler and Freeman, 2017) alongside regions of the extended network and orbito-frontal cortex (Freeman and Johnson, 2016).

Taken together, these findings illustrate the current conceptualisation of the neurocognitive underpinnings of face perception. Currently, the dual-route model (Haxby et al., 2000), with the necessary revisions (Duchaine and Yovel, 2015), seems to have generated a certain degree of consensus among researchers, particularly concerning the functional role of the regions in the core network. Furthermore, researchers in social vision have started to investigate the predictions made by more comprehensive computational models of person perception (Stoler and Freeman, 2016; 2017). However, these studies have exclusively used facial stimuli, possibly because of a higher understanding of the neural processes underlying the visual analysis of the face, and because of the existence of useful perceptual models have provided a theoretical basis for testing further “higher-level” social constructs (Schweineberger and Burton, 2011; Duchaine and Yovel, 2015).

1.3 Social vision: Bodies

The human body, like the human face, is a rich source of social information about the people around us. There is convincing evidence that we extract socially relevant cues from the body even in absence of a face (de Gelder, 2009; Johnson and Shiffrar 2012; Sell et al., 2008; Knoblich et al., 2006; Lawson et al., 2009; Rosental et al., 1979) and when the signals from the face are not reliable (Rice et al., 2013; Aviezer et al., 2012; Noyes et al., 2018). Yet, the perception of these cues from the body has received less attention.

Social-cognitive psychologists have largely investigated how humans extract social categories from the face (Freeman and Johnson, 2016; Cloutier and Macrae,

2007; Macrae et al., 2005), overlooking how we extract social information from the body. Further, in line with the framework of social vision, the importance of building a comprehensive model of person perception has been emphasised (Freeman and Johnson, 2016). A comprehensive person perception model can be achieved by integrating insights from existing face perception models (Haxby et al., 2000; Bruce and Young, 1986) together with knowledge of the processing of social information coming from other modalities (e.g. voices, Yovel and Belin, 2013; motion, Yovel and O'toole, 2016). However, such more comprehensive proposals of person perception models, so far, have strongly relied on what is known about the perception of socially relevant information from the face, overlooking the social signals coming from the body (Freeman and Johnson, 2016; Bagnis et al., 2019; Martin and Macrae, 2010).

In parallel, vision scientists, have remained at some distance from establishing a “standard” perceptual model for the perception of human bodies although some proposals have been advanced (Minnebusch and Daum, 2009; Taylor and Downing, 2011; Urgesi et al., 2007; deGelder et al., 2010), as opposed to the perception of the face. Drawing direct analogies from faces to bodies is likely going to be imperfect due to their fundamental differences in appearance and dynamics. However, the nature of information about others that we can extract from both faces and bodies is highly similar. Thus, several of the experimental approaches that have been used to investigate the extraction of social information from the face can be valuable tools to investigate human body representation, towards the establishment of a perceptual model of human body perception.

Social vision approaches constitute a highly useful framework to build a perceptual model of the human body. Although in absence of facial information we are able to extract identity specific information from the body (Rice et al., 2013; Reed et al., 2006), it is likely that identity non-specific information (e.g. social-categories and emotional state) will reveal to be crucial for establishing a perceptual model of human body processing. In face perception, while there is evidence that identity recognition is heavily reliant on the extraction of configural information (Rhodes et al., 1993; Maurer et al., 2002), it has been shown that the efficiency of social categorisation relies, at least in part, on the processing of a single or a small number of isolated cues (Cloutier et al., 2005). For example, the accuracy of sex judgments is spared after typical manipulations that alter the ability to extract configural information from the face such as inversion or

blurring (Cloutier and Macrae, 2007). Conversely, the spatial configuration of body parts is fundamentally different from faces and there is evidence that body recognition relies on a differential typology, or possibly a less pronounced holistic processing (Harris et al., 2016; Yovel et al., 2010; Brandman and Yovel, 2010; Soria-Bauser et al., 2011; Robbins and Coltheart 2012). Therefore, understanding the extraction of social cues (e.g. sex, age, race) from shape features of the human body stimuli might be the key to elucidate the fundamental perceptual dimensions of human body processing for the establishment of a consensus model of body perception. Further, understanding the processing of social information from the body might be of vital importance for our understanding of how we form impressions about others in general.

In this spirit, the following section will present the state of the art of research on human body perception. The focus of this review and of the empirical work presented throughout this thesis, will concern the perception of the human body and the social signals (e.g. social categories, emotional states) we can extract from its visual appearance.

1.3.1 Body perception: Behavioural evidence

Human bodies, as faces, have a privileged status during stimulus processing. Task-irrelevant human body stimuli (in the form of silhouette or stick figures) capture attention more than other objects (scrambled bodies, object silhouettes, object stick figures and scrambled objects) when attention is occupied by another task (Downing et al., 2004). Human headless bodies also break into awareness earlier than other objects under continuous flash suppression (Stein et al., 2012). Further, bodies in a natural scene are detected as quickly as faces and thus contribute almost equally in person detection (Bindemann et al., 2010). Initial developmental studies suggested that a preference for the typical human body configuration appeared only during the second year of life (Slaughter et al., 2004; Slaughter et al., 2002). This contrasts with faces, which is suggested to develop earlier (Mondloch et al., 1999). However, more recent behavioural work suggests that infants of 3.5 months already have a preference for upright (and not inverted) bodies which maintain first-order relations and proportions among body parts (Zieber et al., 2014; Zieber et al., 2010; Heron-delaney et al., 2011; Hock et al., 2015). Hence, as faces, bodies are powerful social stimuli able to capture

attention in a privileged manner with such a preference appearing within the first year of life.

Efficient perception of the human body depends on the specific configuration of its parts. Manipulations known to disproportionally disrupt normal face processing, such as inversion, or composite effects, also disrupt normal human body processing (Reed et al., 2003; 2006; Robbins and Coltheart, 2012; Willems et al., 2014). Same-different judgments of body postures, but also identity judgments of learned bodies, are faster and more accurate when body figures are presented upright than inverted (Reed et al., 2003; Robbins and Coltheart, 2012). This inversion effect is present with bodies and faces, but not with scrambled bodies and other objects (e.g. houses or scrambled houses), and is reduced for bodies depicting impossible postures (Reed et al., 2006). Interestingly, inversion effects are not present for body parts presented in isolation or for body halves (and faces) divided on the horizontal axis but still present for body halves (and faces) divided on the vertical axis, possibly because these stimuli, being symmetric on the vertical axis, preserve a larger amount of configural information. Strikingly, the body inversion effect seems to heavily rely on the presence of the head, and is not reduced when other body parts (i.e. a leg or both arms) are removed, highlighting an important role of the head information for configural body processing (Yovel et al., 2010; Minnebusch et al., 2010, but see Robbins and Coltheart 2012).

Composite body effects, where holistic processing is shown to result in easier and faster judgements on single halves of a body when the halves are misaligned, have also been reported (Willems et al., 2014; Robbins and Coltheart 2012b). Differently from faces, composite body effects have been reported more convincingly when bodies were misaligned on the vertical axis (Robbins and Coltheart, 2012b; Willems and Verfaillie, 2014) than on the horizontal axis (Soria-bauser et al., 2011). However, it is possible the strength of this effect is also modulated by task demands. For example, composite body effects were found with horizontal body halves in a posture task (Willems et al., 2014) but not on a same-different identity task (Soria-bauser et al., 2011; Robbins and Coltheart, 2012). Evidence for holistic processing of human bodies has been more recently reported by demonstrating a whole-versus-part superiority effect (Harris et al., 2016). Body parts are recognised better in the context of a whole body than alone even when a depth manipulation makes holistic processing of a whole-body figure more difficult (Harris et al., 2016). Taken together, these findings show evidence for

configural-holistic processing of human body stimuli. Furthermore, the existence of mixed results suggests that drawing complete analogies with face processing might be inaccurate. Face stimuli are highly constrained in terms of their spatial configuration, and a distinction between a whole and its functionally significant parts is well defined (Maurer et al., 2002). Human body holistic processing might have a finer resolution: body parts (e.g. an arm) are, in turn, constituted by functionally significant sub-parts (hands, fingers) (Reed et al., 2006; Harris et al., 2016). Further research is needed to clarify whether such effects are modulated by task demands (Willems et al., 2014), and if, differently from faces, are found also with ensembles of body parts (Harris et al., 2016).

In line with the proposal of a face space, an analogous body shape space has been outlined (Hill et al., 2016; Hu et al., 2018; Rhodes et al., 2013; Sturman et al., 2017). Computational approaches have demonstrated that the dimensions of a language-based similarity space of body descriptions closely matched those obtained from a geometric body-shape space generated from three-dimensional laser scans of thousands of bodies (Hill et al., 2016). Given the tight link with language, the authors tentatively interpret the five dimensions of the two spaces as representing: 1) weight, 2) masculine vs curvy, 3) height, 4) waist height, and 5) classically feminine shapes versus other shapes.

These theoretical dimensions outlined by Hill et al. (2016) are in agreement, at least in part, with results coming from visual adaptation studies (Lawson et al., 2009; Sturman et al., 2017; Rhodes et al., 2013; Sekunova et al., 2013; Glauert et al., 2008; Hummel et al., 2012; Brooks et al., 2016; 2019; Challinor et al., 2017; Winkler and Rhodes, 2010; Kovacs et al., 2005). Initial studies have suggested that human body orientation (on the horizontal axis) is susceptible to adaptation effects and that it is coded responding to a multiple-channel model rather than an opponent coding model (Lawson et al., 2009). Here, rather than being referenced to a “normal” frontally oriented body (opponent coding), left-right, and direct body orientations appeared to be coded by three different channels independent from each other (multiple channel model).

Conversely, higher-level human body dimensions, as demonstrated for faces (Leopold et al., 2001; Webster and Macleod, 2011), follow the rules of an opponent coding model. For example, brief adaptation to an identity biases perception to the

opposite of that identity, and the size of the aftereffect increases with the distance between the adaptor and the test stimuli (Rhodes et al., 2013). In support of a multidimensional body space, a recent study has revealed independent adaptation to body fat (in line with previous research, Winkler and Rhodes, 2010; Stephen et al., 2016; Hummel et al., 2012), and body muscle (Sturman et al., 2017). That is, exposure to bodies with extreme fat levels causes fat aftereffects but not muscle aftereffects, whereas exposure to bodies with extreme muscle levels causes muscle aftereffects but not fat aftereffects.

Bodies' sex, another crucial body shape-contingent social dimension, also generates aftereffects when body silhouettes varying in their sexually dimorphic global shape are adapted (Palumbo, Laeng and Tommasi, 2012). Importantly, high-level aftereffects, such as those for body size, occur also across viewpoint, pose and regardless of variation in the low-level properties of the images (Sekunova et al., 2013; Brooks et al., 2018). Taken together, these findings support the existence of a mental space for the encoding of body shapes. Further, adaptation paradigms can inform the cognitive architecture of human body perception by identifying the crucial, neurally dissociable, perceptual dimensions of the human body.

Only recently, the functional independence of human body social dimensions has been assessed using the Garner selective attention task (Johnstone and Downing, 2017; Reed et al., 2018). Garner interference tasks have demonstrated to be informative for confirming and expanding hypothesis brought by existing perceptual models of face perception (Atkinson et al., 2005; LeGal and Bruce, 2002; Schweineberger and Soukoup, 1998;1999). That is, the core predictions of the most influential neurocognitive models of human face perception concern the separability/integrity of processing of different high-level social dimensions conveyed by the face (e.g. Familiarity, Bruce and Young, 1986; Identity and emotional expression/gaze, Haxby et al., 2000). Similarly, Garner approaches can be used with human body stimuli for setting the foundation towards a perceptual model of body perception.

For example, over three experiments Johnstone and Downing (2017) found asymmetric interference between body sex and size using silhouettes and natural body images. That is, the irrelevant variation of sex interferes with body size judgments but

not the converse. In line with previous findings, they interpret the results with a parallel-contingent model: sex influences the ongoing processing of body size. In other words, body shape information relevant for sex processing provides a reference to efficiently process body size. This results in interference, as measured by the Garner paradigm, when sex irrelevantly varies during size judgments. The underlying mechanisms of the asymmetrical Garner Interference pattern can reflect the broad distinction between variable (e.g. size) and invariant (e.g. sex) information that, as for faces (Duchaine and Yovel, 2015), has been proposed to hold for the visual representation of bodies (Vangeneugden et al., 2014; Giese and Poggio, 2003). In line with several Garner studies on face social dimensions of the face (Atkinson et al., 2005; Schweineberger and Soukoup, 1998;1999) asymmetrical interference of invariant information over variant information could be explained by a one-way crosstalk from the static towards the dynamic pathways.

Although largely overlooked, a small number of social-psychological studies have also investigated how we extract social information from the body. For example, human body shape cues are the primary information that people use to decide other's sex (Johnson and Tassinari, 2005). Categorical decisions about other's sex strongly rely on waist-to-hip ratio, a cue that is extracted from the shape of the body, while continuous judgments of gender (i.e. masculinity and femininity) are predicted by motion cues of the body (Johnson and Tassinari, 2005). Further, decisions of sex from visually impoverished stimuli of bodies or body parts are biased towards male (Johnson and Freeman, 2012; Gaetano et al., 2014). Compared to a real distribution of measures of waist-to-hip ratio, people's judgments about others' sex from the body needs a highly "feminine" waist-to-hip ratio, to shift decision from male to female. Similarly, in a recent study, categorical judgments about body emotions were more efficient when compatible with their sex stereotypical association (e.g. female – happy, Bijlstra et al., 2019).

Besides sex, people can also accurately judge body strength from static pictures of human bodies (Sell et al., 2008). These judgements were highly correlated with judgements of formidability (belief about a man's fighting ability) highlighting that the perceived strength is the main cue selected for assessing formidability (Sell et al., 2008). Interestingly, a recent study has found that judgments of physical size and formidability from faces and bodies are biased by race (Wilson et al., 2017). Black male bodies and faces tend to be rated as bigger (taller, heavier and more muscular) and

more physically threatening than the faces and bodies of white men. Altogether, these findings are often interpreted from a distal perspective, in the sense that they reflect high-level biases or adaptive behaviours. However, as current views of dynamic interaction between top-down and bottom up cues propose (Freeman and Ambady, 2011; Freeman and Johnson, 2016), it is possible that they also reflect the perceptual encoding of human body social dimensions.

Taken together these findings show that there is a considerable amount of evidence that implies that human bodies, like faces, have a privileged status during stimulus processing. Such attentional privilege is acquired at early stages of development (Zieber et al., 2010). Moreover, a growing number of adaptation studies is revealing the nature of the perceptual dimensions of human bodies (Lawson et al., 2009; Sturman et al., 2017). These studies have shown that high-level socially relevant dimensions conveyed by the human body are sensitive to perceptual aftereffects (Sturman et al., 2017), supporting the validity as a framework of a mental space for human body shape (Hill et al., 2016). However, the relative contribution of each stimulus dimension during stimulus processing has not yet been explored. Understanding the relationship between dimensions during stimulus processing, for example using Garner approaches, has allowed the confirmation of existing, or the development of new predictions on perceptual models of face perception (Schweineberger and Soukoup, 1998;1999; Atkinson et al., 2005), and constitutes a crucial step towards a model of human body perception.

1.3.2 Body perception in the brain

Human bodies are a rich source of socially relevant information, they are visually salient and can capture attention before other objects (Downing et al., 2004). Their special status during information processing is supported by the considerable cortical resources that the brain devotes to represent the visual information of the human body (Peelen and Downing, 2007; Minnebusch and Daumn, 2009; Berlucchi and Aglioti, 2010). Specifically, occipitotemporal brain systems containing functionally selective brain regions, are thought to underlie the visual representation of the human body form and to carry out computations that give the human body its special status during stimulus processing.

In the nonhuman primate brain, studies have revealed that neurons in the inferior temporal cortex (IT) respond selectively to the appearance of human and monkey bodies and body parts (Pinsk et al., 2005; Kiani et al., 2007; Wachsmuth et al., 1994; Desimone et al., 1984; Gross et al., 1972; Gross, 1992). For example, when viewing a large number of images from natural and artificial object categories, patterns of activity of a large number of neurons (>600) in monkey IT showed a distinct pattern for animate and inanimate objects (Kiani et al., 2007). The category of animate objects was divided into distinct patterns for faces and bodies with further distinction between different types of faces (human, monkey or animal) and different types of bodies (human, four-limbed animal or bird). Beyond the IT, studies using single cell recordings with non-human primates have reported neurons in the anterior part of the superior temporal sulcus (a-STS) of the monkey that were responsive to static body postures, particularly when they were representing the endpoint of an action (Jellema and Perrett, 2003). Functional MRI studies on primates identified regions in the a-STS responding to bodies and body parts (Tsao et al., 2003, Tsao et al., 2008; Pinsk et al., 2005; Popivanov et al., 2012; 2014; Kumar et al., 2019). Interestingly, body selective patches are found also in the macaque homologue of fusiform face gyrus (Tsao et al., 2003; Pinsk et al., 2005; Tsao et al., 2008). However, these patches identified with fMRI, though they show stronger responses for bodies compared to other non-face objects, still have a greater response for faces (Tsao et al., 2008; Popivanov et al., 2012). Such comparatively low category selectivity for bodies in the body patches is likely due to the presence of neurons in those patches that show a high “within-category” selectivity (single neurons responding to certain body stimuli but not others) (Popivanov et al., 2014). This suggests that these neurons respond to visual features that occur frequently, but not exclusively, in human bodies (Popivanov et al., 2016). More recently, a study directly compared fMRI-defined a-STS body patches with the more posterior middle STS body patches (Kumar et al., 2017). Single cell recordings in these patches showed an increasingly viewpoint-independent representation for monkey body posture and identity from posterior to anterior STS, similarly to faces (Freiwald and Tsao, 2010). Together these findings show that non-human primates show body selectivity in temporal cortex. Importantly, fMRI studies on primates can help to draw parallels between human fMRI studies and single cell recordings by capturing the large-scale representation of human body information.

In humans, evidence for body specific representations comes from studies on lesioned patients (Schwoebel and Coslett, 2005; Corradi-Dell'Acqua and Rumiati, 2007). Pick (1922) identified patients who were unable to name body parts on themselves or others (a condition referred to as *autotopoagnosia*). More recently, an autotopoagnosic patient (G.L.) with extensive left hemispheric damage, was found to be impaired in pointing or visually identify body parts on himself or others, and matching same body parts across changes in viewing angle (Buxbaum and Coslett, 2001). Although such neuropsychological disorders such as autotopoagnosia have been rarely reported, often being confounded with more general visuo-spatial impairments (Derenzi and Scotti, 1970), lesions on ventromedial occipitotemporal regions corresponding with body selective regions EBA and FBA identified with fMRI (Downing et al., 2001; Peelen and Downing, 2005), are associated with defective recognition of bodies and body parts (Moro et al., 2008). Patients with occipitotemporal lesions were less accurate in a delayed match-to-sample task involving body parts compared to face and object parts. Further, evidence for dissociable body representations comes from intracranial surface electrode recordings of electrical activity in human brains performed on open scalp epileptic patients (Pourtois et al., 2007). Visually evoked potentials using this technique were highly selective, with an increased response to images of bodies relative to faces, mammals and tools in the right lateral occipitotemporal cortex starting at 190 ms and peaking at 260ms post-image onset.

Perhaps the most compelling evidence of body selectivity in the human brain comes from fMRI studies comparing brain response to images of headless bodies and body parts against responses to objects, object parts, faces, scenes and other control images (Downing et al., 2001; Downing and Peelen, 2011; Peelen and Downing, 2007). The functional data of studies converge to present two strongly body selective, focal regions: the extrastriate body area (EBA) (Downing et al., 2001), located in the posterior inferior temporal sulcus/middle temporal gyrus, and the fusiform body area (FBA) (Peelen and Downing, 2005; Schwarzlose et al., 2005), located in the fusiform gyrus. EBA, depending on the statistical threshold used to localise the area, overlaps with human motion selective area MT (hMT) and with the dorsal focus of the object selective area (LO) (Downing et al., 2007), particularly on group level analyses. However, studies using multi voxel pattern analysis (MVPA) have revealed that these stimuli elicited

independent activity patterns in the lateral occipito-temporal cortex (Downing et al., 2007; Peelen et al., 2006).

The FBA is adjacent and partly overlaps with the fusiform face area (FFA, Kanwisher et al., 1997). Nonetheless, region of interest (ROI) analyses of non-overlapping body- and face- selective voxels in high resolution fMRI show that the FFA ROI did not respond to bodies and the FBA ROI did not respond to faces (Schwarzlose et al., 2005). In line with these data, a study using MVPA shows that local patterns of selectivity for faces and bodies across voxels of the fusiform gyrus are unrelated (Peelen and Downing, 2006; Kaiser et al., 2013). Together, these findings suggest that the occipitotemporal cortex is equipped with regions involved specifically in performing visual analysis of the human body. These regions, with careful localisation within studies and within participants, can be dissociated from overlapping activations coming from other category-selective regions (i.e. faces, objects, and motion).

1.3.2.1 Viewpoint and Identity. The body selectivity of EBA and FBA is not limited to realistic pictures of human bodies, but extends to more abstract representations of bodies such as line drawings, “stick figures”, and silhouettes (Downing et al., 2001; Peelen and Downing, 2005). This distinction indicates that these regions encode body representations that are abstract across specific visual features. While body selective areas can abstract across stimulus features, responses in these regions appear to be viewpoint- dependent (Chan et al., 2004; Taylor et al., 2010; Saxe et al., 2006). That is, changes in view greater than 45 degrees causes a release from adaptation in EBA and FBA (Taylor et al., 2010). Furthermore, bodies and body parts presented in third person view, as opposed to first person view, elicit higher activation of right EBA (Chan et al., 2004; Saxe et al., 2005; 2006). Moreover, the body representations in these regions are stronger when bodies appear in commonly experienced configurations (right part of the body in the observer’s left visual field) in support of a preference for allocentric views that reflect the statistical occurrence of bodies in context (Chan et al., 2010). Together, these findings support that EBA and FBA respond to high-level visual features that distinguish a body from other objects, as they are sensitive to abstract representations of the bodies, and play a role in parsing the depicted shape of a body (due to their sensitivity to viewpoint), particularly when seen in third person perspective.

Viewpoint effects, and a stronger activation for allocentric body views (at least in the right hemisphere), advance the possibility that these regions are involved in the detection of other people. Another question relates to whether EBA and FBA representations are directly involved in body recognition, i.e. in discriminating bodies of different identities. Several studies have investigated body identity, for example, by testing if EBA and FBA differentially respond to own vs another (familiar or unfamiliar) person's body (Chan et al., 2004; Hodzic et al., 2009a; Devue et al., 2007; Vocks et al., 2010). For example, Chan et al. (2004) in a blocked design fMRI experiment report no effect of the self vs familiar others in EBA or FBA. Similarly, comparing the fMRI activation evoked by own versus others familiar bodies, Hodzic, et al. (2009) found no difference in EBA or FBA. Instead, the contrast analyses of Hodzic and colleagues (2009) identified parietal regions in the right hemisphere (right inferior parietal lobule, right inferior parietal sulcus) and frontal regions in the left hemisphere (left posterior orbital gyrus). Together, these findings suggest that these regions are not directly involved in distinguishing own versus other bodies. Mixed findings have been instead reported when the view of own or familiar bodies was compared with unfamiliar others (Hodzic et al., 2009a; Vocks et al., 2010; Hodzic et al., 2009b; Cazzato et al., 2015). In all cases, when effects in occipitotemporal regions were found (Hodzic, Kaas et al., 2009a; Vocks et al., 2010), the data are in the direction of higher activation in EBA and FBA for own vs others unfamiliar bodies.

Other studies used fMRI adaptation techniques (a-fMRI) to test whether a release from adaptation was present in EBA or FBA when changing body identities (Ramsey and Hamilton, 2010; Wiggett and Downing, 2011; Myers and Sowden, 2008; Kable and Chatterjee, 2006). For example, presenting videos of whole body and face-stimuli with repeated identity or action elicited adaptation for repeated actions but not for repeated identities in occipito-temporal cortex (Kable and Chatterjee, 2006; Wiggett and Downing, 2011, but see Ramsey and Hamilton, 2010). These findings suggest that body selective regions EBA and FBA have little or no role in processing the identity information from the body, especially beyond more consistent effects of viewpoint (Chan et al., 2004; Taylor et al., 2009). This supports the idea that these regions distinguish a rather more general difference in body shape and posture (Downing and Peelen, 2011). Importantly, the body information computed in these regions is likely to be used by other regions to differentiate between identities. For example, recent and converging findings

in face, body or whole person perception studies, propose that body and face form information are progressively integrated in the anterior temporal lobes (Harry et al., 2016; Rajimehr et al., 2009; Taylor et al., 2007; Taylor and Downing, 2011; Tsao et al., 2008; Fisher and Freiwald 2015; Greven and Ramsey, 2017). In line with recent findings in macaques fMRI studies using face stimuli (Chan and Tsao, 2017), it is possible that these higher level, integrated representations in the anterior temporal lobes are a key factor to differentiate person's identity.

1.3.2.2 Body in motion and in action. Together with systems specialised for selectively processing facial and body form, our brains contain systems specialised for processing movements performed by bodies and faces. Most of the studies investigating the perception of human motion from the body have used point-light-displays (PLD), which can convey the pattern of biological motion with minimal or absent body form information. Functional imaging studies have used PLD stimuli to compare brain activity in response to biological vs non-biological human body motion (Bonda et al., 1996; Grossman et al., 2000; Grossman and Blake, 2001; Grossman and Blake 2002; Peuskens et al., 2005), and identified a portion of the posterior superior temporal sulcus (pSTS) that selectively responds to human movements. Further, selectivity for biological (whole-body) motion has also been reported in the inferior temporal sulcus/ middle temporal gyrus (Grossman and Blake 2002; Peuskens et al., 2005). Whole-body biological motion also elicits a significant response in functionally localised body- and face- selective regions (Grossman and Blake 2002; Peelen and Downing, 2006). However, the degree of selectivity for intact vs scrambled PLDs in body-selective regions correlated positively with the degree of selectivity for static body images (and not with motion or face selectivity) (Peelen and Downing, 2006). Thus, it is likely that body selective regions EBA and FBA respond to biological motion stimuli in virtue of a reconstruction of the body form from the motion pattern (structure-from-motion) rather than being directly involved in processing body motion (Peelen and Downing, 2006).

The perception of body dynamics is an initial step for the perception of other's actions (Hamilton and Grafton, 2007). The perception of body actions activates multiple regions in the posterior temporal cortex including the pSTS and the EBA (Grosbras et al., 2012). One possibility that could account for this finding, is that body selective regions contain representations selective for dynamic body actions and/or for static

bodies implying an action (Lange and Lappe, 2006; Jarstorff and Orban, 2009). For example, typical and atypical grasping movements towards objects elicited higher activation of functionally defined EBA compared to a control condition where objects were being touched (Valyear and Culham, 2010). However, transcranial magnetic stimulation (TMS) over EBA impairs discrimination of body form but not of coherent body postures implying an action (Urgesi et al., 2007). Discrimination of body actions was instead impaired when stimulation was delivered to ventral premotor cortex (vPM) (Urgesi et al., 2007; Candidi et al., 2008). In line with these data, presentation of a coherent series of frames depicting an action elicited higher activation in the pSTS, while the same actions presented in a series of incoherent frames (thus generating an incoherent sequence of movements) more strongly activated the EBA (Downing et al., 2006).

In a more recent study, Vangeneugden et al. (2014) provided converging evidence of a double dissociation between neural mechanisms of body form and body motion discrimination using fMRI, psychophysical approaches and TMS methods. In three experiments they presented point-light walkers varying on two orthogonal dimensions: the facing orientation of the walker (leftward or rightward), and the walking direction of the walker (forward or backward). The former is related to the dimension of form and the latter to the dimension of motion. Patterns of activity within functionally localised EBA (obtained by subtracting correlations of conditions with same vs different facing orientation) discriminated the facing orientation but not the walking direction. Conversely, patterns of functionally localised STS (obtained by subtracting correlations of conditions having the same vs different walking direction) discriminated walking direction but not facing orientation. In a further psychophysical experiment the authors manipulated point light display in two ways: misaligning the ellipses of a point light walker or reducing the presentation duration of the walker. The former manipulation affected behavioural performance in a form-related facing orientation task but not on the motion-related walking direction task, the latter manipulation affected performance on the walking direction task but not the facing orientation task. In a third experiment, offline repetitive TMS over functionally localised STS disrupted performance in the walking direction task but not in the facing orientation task, while TMS over EBA affected performance on the facing orientation task but not in the walking direction task.

Taken together, these findings strongly support a dissociation between the processing of biological body motion (representing both transitive or intransitive actions) in the pSTS and the processing of the body form in EBA/FBA. This view is in line with biological motion perception models that posit two separate pathways: one ventral, analysing form signals, and one dorsal, analysing motion signals (Giese and Poggio, 2003).

1.3.2.3 Body emotion. Characteristic body postures and body movements can also convey basic emotions (Atkinson et al., 2004; DeGelder, 2009; de Meijer, 1989; Dittrich et al., 1996; Wallbott, 1998). How do body form and body motion selective regions respond to bodies conveying emotional information? Findings from face perception suggest that dynamic and static emotional faces modulate responses in visual cortex, particularly in the FFA. These responses are contingent on the activation of the amygdala (Vuilleumier et al., 2004; Vuilleumier, 2005; Vuilleumier and Pourtois, 2007; Vuilleumier et al., 2001). Similarly, static and dynamic emotional bodies or body parts modulate responses in visual cortex (Hadjikhani and de Gelder, 2003; Grezes et al., 2007; Peelen et al., 2007; Grosbras and Paus, 2005). Specifically, such modulations have been reported in the fusiform gyrus, presumably corresponding to the FBA (Peelen et al., 2007 (functionally localised); Van de riet et al., 2009; Grosbras and Paus, 2005) in the lateral occipitotemporal cortex (Peelen et al., 2007 (functionally localised); Grezes et al., 2007; but see Van de riet et al., 2009) and in bilateral pSTS (Grezes et al., 2007; Pichon et al., 2008), particularly with dynamic stimuli expressing fear (Grezes et al., 2007). Using MVPA, a study has demonstrated that the increased response to emotional body expressions in body selective regions EBA and FBA was related, on a voxel by voxel basis, with the degree of selectivity for static bodies (and not with the selectivity for faces), in support of a specific enhancement of body selective neural populations as a function of the expressed body emotions (Peelen et al., 2007). Further, these responses to emotional stimuli in visual cortex are positively correlated with the amygdala response (Peelen et al., 2007; Hadjikhani and de Gelder, 2003; Hortensius et al., 2017), although to a lesser extent than faces (Kret et al., 2011) and sometimes even in the absence of an emotional expression, with neutral dynamic body stimuli (Grezes et al., 2007).

Importantly, it is possible that the emotional modulation of visual responses is partly or entirely driven by the salience (arousal) of the emotional stimuli rather than by

their emotional valence. Indeed, several studies have been using highly arousing emotions such as angry, fearful or happy body expressions (Grezes et al., 2007; Hadjikani and de Gelder, 2003; Pichon et al., 2008). When sad emotional expressions were used no emotional modulation of body selective visual regions was found (Peelen et al., 2007). However, the response of body selective visual regions does not seem to be modulated by other arousing stimuli related to the body, as the perception of others' pain (Decety and Lamm 2006; Morrison et al., 2007; but see Gu et al., 2010). Together, these findings suggest that emotional expressions modulate response of regions specialised for processing bodies (and faces) perhaps to allow a more efficient discrimination of other's emotional states. Moreover, a study has suggested that the pSTS might have a wider role than increased sensory representation of motion during emotion perception (Peelen et al., 2010). The pSTS (and mPFC) showed similar patterns of activity across stimulus modality (emotional bodies, faces, or voices) but distinct patterns for the emotion categories tested (happiness, sadness, anger, disgust, fear).

1.3.2.4 Extracting social categories from the body. Human bodies also signal cues about social categories such as others' sex, health, weight, and personality (Hu et al., 2018; de Gelder, 2009; Johnson and Tassinari, 2005). What are the neural mechanisms by which these signals are detected? A recent study investigated how body-responsive brain regions encode information about the subcategories of sex and weight (Foster et al., 2019). Using MVPA they found evidence for higher than chance decoding of body sex in EBA and FBA. Surprisingly, significant sex decoding was also evidenced in face selective ROIs OFA and FFA and V1. Similarly, higher than chance decoding of body weight was observed in all these ROIs, but the FBA. Importantly, when controlling for the size of the stimuli, only EBA and FBA showed a significant response for sex and weight, while face selective regions (except FFA for weight) and V1 did not. This finding is in line with recent behavioural evidence that shows partial-integral processing of these two dimensions in a visual task (Johnstone and Downing, 2017).

Activity of occipitotemporal body selective regions has been reported in a study investigating social stereotypes (Quadflieg et al., 2011) and the formation of impressions (Greven et al., 2016; 2018; Greven and Ramsey, 2017). Using whole person stimuli, Quadflieg et al. (2011) found that functionally localised EBA and FBA

showed higher univariate response to pictures of people that, based on their clothing, violated stereotype consistent expectations (e.g. a male nurse). Beyond visual regions, this response was found to functionally interact with response in dorsolateral prefrontal cortex (dlPFC), a region involved in conflict monitoring (Quadflieg et al., 2011). In line with a sensitivity of body perceptual regions to social information, another study showed increased response in right FBA, and bilateral EBA when contrasting muscular to slim bodies but not for obese over slim bodies. Analyses of functional connectivity showed that the contrast obese over slim bodies led to increased coupling between EBA and the theory of mind network and the temporal poles (Greven, Downing and Ramsey, 2018). Similar studies have found functional interplay between body perception regions and theory of mind network when forming an impression linked to a description or recalling an impression previously formed about others' bodies (Greven et al., 2016; Greven and Ramsey, 2017).

These findings show that body selective regions and the connections between these regions and other networks relevant for social perception, are sensitive to social categorisation. Moreover, they are in line with recent models of person perception which posit a role for visual regions in detecting crucial information relevant for social categorisation (Freeman and Johnson, 2016). Importantly, it is possible that task-relevance, and thus attention to those dimensions relevant for categorization, is needed for these activations to be observed. Indeed, increased activity in visual cortex was not observed when participants were not explicitly asked to categorise sex (Quadflieg et al., 2011) or when they were not explicitly asked to form impression on bodies (Greven et al., 2018). To date, the only study that has reported that social dimensions elicit activity in body selective cortex in a task-irrelevant manner is the most recent by Foster and colleagues (2019).

1.3.2.5 Current models of human body perception. How do these findings, related to the functional specificity and functional properties of body selective regions, contribute to a theoretical model for the perception of the human body? A possible approach for establishing a perceptual model of human body perception may be to draw parallels with existing models of human face perception (Minnebusch and Daum, 2009; Taylor et al., 2007) or of other category-selective regions (Taylor and Downing, 2011). Indeed, beyond the anatomical proximity of face- and body- selective regions, bodies and faces

are nearly always perceived together in space and time, and both carry similar social information.

Influential models of face perception identify a core and an extended system for face processing (Haxby et al., 2000; Duchaine and Yovel, 2015). The core system carries out the visual analysis of the face, the extended system is not directly involved with the visual analysis of the face but extracts information from regions in the core systems. Generally, within the core system, these models identify a posterior-to-anterior increasingly complex face-specific representation (from face parts to the whole face, to identity representations). In parallel, they also posit a variable/dynamic vs invariant/static distinction expressed by a ventral and a dorsal system (Hoffman and Haxby, 2000; Duchaine and Yovel, 2015).

Several findings from fMRI confirm similar distinctions in body-selective regions. Within the static pathway, Taylor et al. (2007), drawing an analogy between the functional properties of OFA/EBA and FFA/FBA (Yovel and Kanwisher, 2005; Harris and Aguirre, 2008; 2010) tested whether a part versus whole body distinction existed between EBA and FBA. Across three fMRI experiments, when the amount of body information displayed was varied gradually from a single body part to a whole body, EBA showed a linear increase in proportion to the amount of the body that was visible (the more body parts were visible the higher activation of the region). In contrast, FBA showed a quadratic increase in response to torsos and headless bodies relative to images of smaller body parts, in that its activity steeply increased when whole torsos were displayed. This suggests that EBA carries a relatively part-based representation of the body, while FBA is more sensitive to second-order relationships among body parts.

Two other models of body processing have proposed that EBA is preferentially involved in the representation of body parts. In one case, based on evidence from transcranial magnetic stimulation, Urgesi et al. (2007) advanced the hypothesis that body processing follows two routes: one “ventral” route including EBA, which processes local features of bodies as body parts and body form, and one “dorsal” route, including vPMC, which processes the whole body in a configural manner making use of sensory-motor representations. In another model, Hodzic et al. (2009) proposed the existence of distinct networks for body detection and body identification. In their model they assume that EBA is particularly involved in the detection of bodies and body parts, while FBA

was sensitive to body recognition showing a higher response for familiar than unfamiliar bodies. Moreover, a recent paper has suggested that responses in the anterior temporal lobes to whole-face and whole-body were positively correlated on a voxel by voxel basis (Harry et al., 2016). This supports an increasingly complex posterior-to-anterior body representation and, together, a convergence of distinct sources of perceptual information related to the whole person in the anterior temporal cortex.

In support of the ventral/static and dorsal/dynamic distinction adopted from face models (Hoffman and Haxby, 2000), several findings have convincingly reported a dissociation between regions involved in analysing the body form and the body motion information (Peelen and Downing, 2006; Vangdeneugden et al., 2014). EBA and FFA would constitute the static pathway processing static information about the body form while the pSTS would be part of a dynamic pathway analysing information about body motion. Such a distinction has also been proposed in influential models of action perception (Giese and Poggio, 2003).

In support of the existence of an extended network for body perception, there are at least two lines of evidence for regions indirectly involved in supporting the activity of the “core regions” for the visual analysis of body perception. First, there is consistent evidence that the perception of body emotions, particularly the arousing emotions, is mediated by the activation of the amygdala (Hadjikhani and de Gelder, 2003; Grezes et al., 2007; Peelen et al., 2007; Pichon et al., 2008). Second, studies of functional connectivity have highlighted that body-selective regions work in concert with other social perception systems (e.g. Theory of mind network, executive functions), when forming impressions or recalling memories about others (Ramsey, 2018; Greven et al., 2016; 2018; Greven and Ramsey, 2017a,b) or when a stereotype is activated (Quadflieg et al., 2011).

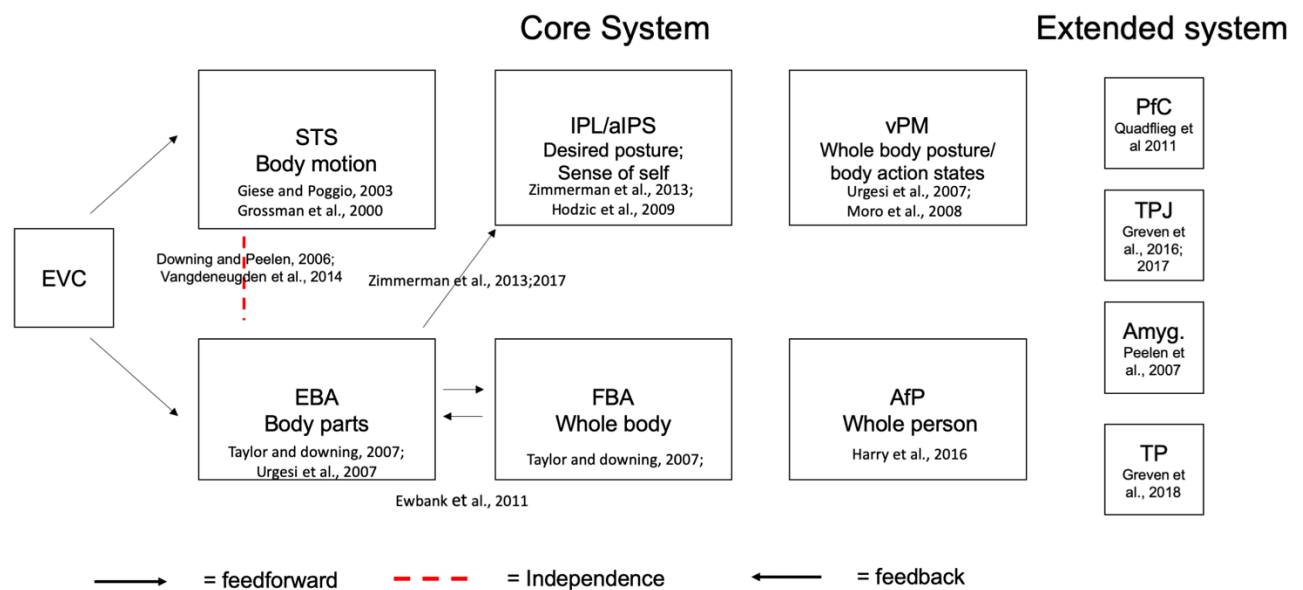


Figure 1.1 Parallel between Haxby model and findings in human body perception literature.

Together, this evidence suggests that the logic applied in existing face models could be beneficial to understand the functional properties of body selective regions (Figure 1.1; Minnebusch and Daumn, 2009; Taylor et al., 2007, but see deGelder et al., 2010). However, although several functional claims have been made for regions that show a body specific response, research on body perception is still far from a standard perceptual model. Indeed, particularly for EBA and FBA, several functional properties have been claimed: from a “simple” generation of a body percept (Downing and Peelen, 2011), to self/other distinction (Hodzic et al., 2009), or even motor control in absence of vision (Astafiev et al., 2004).

The disproportionate amount of research in face perception across several fields of psychology and neuroscience (Little et al., 2011), and the solid neurocognitive frameworks that have been proposed (Haxby et al., 2000; Bruce and Young, 1986), has progressed our understanding of the perception of the face beyond the solely cognitive aspects of face processing (e.g. part- versus whole-based processing) (Collishaw and Hole, 2000; Tanaka and Farah, 1993; Liu et al., 2002; Hoffman and Haxby, 2000). Indeed, in the last decade, following the program of social vision, researchers have focussed also on how socially-relevant cues, extracted from the appearance of the face,

support the inferences people make about others (Adams et al., 2011; Johnson & Shiffrar, 2012; Fiske, Gilbert, & Lindzey, 2010).

Several studies in face perception have confirmed the validity of social vision approaches, and have led to the proposal of face perception models under this framework (Freeman and Ambady, 2011; Adams and Kveraga, 2015; Freeman and Johnson, 2016). Recently, neurocognitive research has shown that socially relevant information from the face shapes perceptual representations of the face in visual regions (Stolier and Freeman, 2016; Quadflieg et al., 2011; Stolier and Freeman, 2017).

Our knowledge of the perception of the human body is somewhat advanced at the level of its implementation. As summarised in Figure 1.1, several neuroimaging studies have identified where body perceptual processes are implemented in the brain (Downing and Peelen, 2007; Peelen and Downing, 2011). Several of these studies have also directly investigated the functional properties of these regions (Taylor and Downing, 2007; Vangeneugden et al., 2014; Peelen and Downing, 2006; Harry et al., 2016; Peelen and Downing, 2011). However, our understanding of the cognitive architecture of human body processing from a functional point of view has not yet reached a unified framework.

The empirical work of this thesis focuses on understanding the computations behind human body perception from a functional-psychological point of view and, at this stage, gives less attention to its implementation. In doing so, it aims at building that ground cognitive work that may allow for the development of a consensus in body perceptual models to arise. What are the relevant dimensions that we extract from the human body? How are such dimensions related to each other? Establishing the presence or absence of fundamental interactions between the multiple, relevant information conveyed by human body can, in turn, generate precise predictions on the implementation of these information in the brain.

Further, in line with novel face perception frameworks proposed under the program of social vision (Freeman and Johnson, 2016), the following empirical chapters particularly focus on the extraction of socially relevant information from the appearance of the human body. The utility of adopting this approach for body perception is twofold: on one hand, it can shed a light on the cognitive architecture of the crucial social dimensions conveyed by human body stimuli, thus expanding the existing proposal of

unified frameworks of body perception. These frameworks have largely overlooked the social cues extracted from the body (Minnebusch and Daumn, 2009; Taylor and Downing, 2007); on the other, they will investigate the provocative proposal of face models formed under social vision approaches, and the notion that initial percepts (e.g. the shape of a body) might be highly integrated with the social information they unequivocally convey.

1.4 Overview of the thesis

In the current chapter (Chapter 1), the historical background that led Social Vision approaches to arise has been introduced. Further, following discussion of the logic of this approach and an example of person perception model under this framework, the relevant behavioural and brain imaging literature on face and body perception have been reviewed. The findings reported mainly aimed to cover two main issues: 1) Overview of findings that have informed existing perceptual models of the perception of social stimuli; 2) Overview of findings that investigated how humans extract social categories from faces and bodies.

Chapter 2. Investigates the perceptual representation of body sex. While it is known that body sex is an important perceptual dimension conveyed by the body (Hill et al., 2016; Palumbo et al., 2012) this is the first set of studies that investigate how, within the sex dimension, male and female body perceptual representations are organised. In six experiments, using a visual search paradigm and the logic of search asymmetry, the perceptual encoding of body sex is unravelled.

Chapter 3. Investigates the relationship between the processing of body emotion and body sex. Across three experiments, the relationship between two socially relevant body dimensions is explored using the Garner selective attention paradigm (Garner, 1974). In line with previous literature in face perception, such paradigms directly contribute to establish perceptual models.

Chapter 4. Investigates the neural mechanisms by which long-term expectations based on a body characteristic (i.e. sex) are expressed. In doing so, this chapter explores the possibility that socially relevant characteristics, or their association with specific body shape topologies, is integrated with body perceptual representations formed in the

extrastriate cortex. Further, the neural mechanisms of perceptual expectations towards visual categories represented in ventral visual cortex in general is examined, by investigating how scene-related expectations are expressed with pre-activation of scene-selective regions.

Chapter 5. Summarises the findings from the empirical chapters (chapters 2, 3, and 4). Discusses the perceptual encoding of socially-relevant information from the human body. Specifically, it discusses how male and female body representation are related to each other and how the dimension of sex relates to body emotional expression. Furthermore, the chapter reports how the empirical results contribute to a perceptual model of the human body, and to a better understanding of the perceptual operations during rapid social categorisation.

CHAPTER 2

Visual perception of sex from human body shape is asymmetrical

This chapter investigates the perceptual encoding of sex from human body shape using the logic of visual search asymmetries. Here, we implement social vision approaches by exploring the perceptual encoding of the body within the social dimension of sex. In doing so, we investigate how male and female body shapes are related to each other across 6 behavioural experiments.

This chapter is currently from the manuscript in preparation:

Gandolfo, M. and Downing, P.E. Visual perception of sex from human body shape is asymmetrical.

Acknowledgements:

We are grateful to Kami Koldewyn, Marius Peelen, and Richard Ramsey for helpful comments on earlier versions of the manuscript. Dominic Blackmore-Beales contributed to data collection for Experiment 3.

2.1 Abstract

We efficiently infer others' states and traits from their appearance, and these inferences shape our social behaviour. One key trait is sex, which is strongly cued by body shape. We investigated the structure of visual body shape representations in relation to sex. Specifically we tested a hypothesis, drawn from previous findings in sex-discrimination tasks, that the mental encoding of male and female body shapes is asymmetrical, such that female bodies are positively coded with reference to a male default. If so, following Treisman's search asymmetry logic, then in visual search female bodies should be found more efficiently among male distractors than vice versa. This pattern was confirmed, in both male and female participants, for body silhouettes seen from side and frontal views, and for highly schematised icon figures. Further experiments showed that the search asymmetry favouring female body shape emerged from holistic rather than part-based perceptual processes, and was not explained by lower level stimulus properties. These findings demonstrate an organising principle of the encoding of body shape that supports inferences about a socially relevant characteristic. We speculate on the nature of the features implicated in this search asymmetry, and offer possible accounts for how it arises in the first place.

Keywords: social vision; body perception; sex perception; visual search; search asymmetry

2.2 Introduction

From just a glimpse of another person, we make inferences about their current states and longstanding traits. These inferences are normally spontaneous and effortless; they are also consequential, in that they shape our social cognition and behaviours. Accordingly, it is vital that we understand how such inferences are made (Adams et al., 2011; Johnson and Shiffrar, 2012; Fiske et al., 2010). While the major focus in this area has been on visual face perception (Bruce and Young, 2013; Duchaine and Yovel, 2015; Haxby et al., 2000), we also know that the visual appearance of the rest of the body provides a rich source of socially relevant cues to others' emotions, goals, health, age, strength, and – the focus of the current investigation – their sex (Aviezer et al., 2012; de Gelder, 2009; Rice et al., 2013; Rosenthal et al., 1979; Sell et al., 2009; Shiffrar et al., 2006). Judgments of sex¹ from the physical appearance of others' bodies are generally immediate and highly efficient, even for impoverished stimuli (Gaetano et al., 2014; Johnson and Tassinari, 2005; Kozlowski and Cutting, 1977; Palumbo et al., 2013). What are the underlying mental representations of body shape that support this ability?

In line with earlier work positing a mental face space consisting of multiple dimensions that each capture aspects of facial appearance (Chang and Tsao, 2017; Leopold et al., 2001; Oosterhof and Todorov, 2008; Turk and Pentland, 1991), an analogous body shape space has been proposed (Hill et al., 2016; Sturman et al., 2017). Sex is likely to constitute a key dimension of this space given its behavioural relevance, and in line with recent psychophysical studies. For example, extended viewing of an adapter body of one sex produces a contrast after-effect, such that a briefly presented ambiguous test body is more likely to be seen as being of the opposite sex (Palumbo et al., 2013). In line with other findings of high-level adaptation for social stimuli such as faces (Webster and MacLeod, 2011) and voices (Schweinberger et al., 2008), results such as these are taken to indicate that visual body appearance is represented over a continuous dimension that relates sex to aspects of body shape.

The mental mapping of sex to body shape may not be veridical. Specifically, investigations of visual sex categorization from the body have reported a male bias

¹ In line with current guidance (e.g. <https://orwh.od.nih.gov/sex-gender>), we use the term “sex” to refer to a categorical distinction between simplified body shapes as instantiated in the present experiments, and “gender” elsewhere (e.g. in reference to our participants).

(Gaetano et al., 2012; Johnson et al., 2012). One manifestation of such a male bias is found in the influence of waist-to-hip ratios on judgments of sex from body shape. Relative to a large ground-truth distribution of measured ratios, observers' judgments are shifted, such that some ratios that are generally female in the population are instead frequently reported by observers as male in a forced-choice task (Johnson et al., 2012). Furthermore, observers judge images that are ambiguous by design – such as bodies or hands that fall on the midpoint of a morph between male and female anchor images – as being “male” (Gaetano et al., 2016).

Such findings can be interpreted from a distal perspective, in terms of their potential functional relevance for an observer, and also from the standpoint of proximate mental mechanisms. Regarding the former, researchers have suggested the hypothesis (which we do not address here) that a male bias may be valuable in that males, relative to females, are more likely to present a physical threat. In other words, to assume male incorrectly may be less consequential for the observer than to assume female incorrectly (Haselton and Buss, 2000; Haselton and Nettle, 2006). Considered from the proximate perspective, our focus here, a male bias suggests that for the processes that support perceiving others' sex, the default output is a “male” percept – “female” is determined only in the presence of additional, positive perceptual evidence. Where clear sensory signals indicating “female” are not definitively present, then the perceptual system will settle on a “male” interpretation.

On this perspective, the mental representations of male and female body shapes are not symmetric with respect to each other. That is, while male and female body shapes share many properties in common (e.g. basic structure of trunk and limbs) the female shape is encoded by reference to an additional feature or features, relative to the male. In this respect, the logic resembles a coding scheme offered by Treisman to describe our representation of visual dimensions more generally (Treisman and Souther, 1985; Treisman and Gormican, 1988; Wolfe, 2001). That view argues that coding of elementary visual properties such as colour, closure, orientation, etc. is organised around canonical values and deviations from those values. This account is supported, for example, by findings that visual search for a deviant target shape (e.g. an ellipse) or colour (e.g. lime) among distractors comprised of their canonical counterparts (e.g. a circle, or green) is more efficient than searching for the canonical stimuli among distractors that deviate from them. The interpretation is that departures from a reference

value are positively coded via the presence of additional features. Because presences are more readily detected than absences (Neisser, 1963), this results in more efficient search when the target is a deviant and the distractors take the canonical form than *vice versa*.

Many studies of visual search in general, and search asymmetries specifically, have been applied to discover the representation of relatively simple or “low level” visual properties. Yet search has also been used extensively to understand the “higher level” representations of faces and objects, and the asymmetry logic has likewise been applied to understanding the encoding of more complex emergent stimulus properties (Becker et al., 2011; Cohen et al., 2016; Enns & Rensink, 1990; Enns and Rensink, 1991; Hulleman et al., 2000; Kristjánsson and Tse, 2001; Levin, 2000; Ro et al., 2007; Sun and Perona, 1996a; Sun and Perona, 1996b; Suzuki and Cavanagh, 1995; Tong and Nakayama, 1999). However, studies of visual search for bodies are relatively uncommon (Bindemann et al., 2010; Reeder et al., 2015; Ro et al., 2007), and provide an ideal opportunity to test the asymmetry hypothesis outlined here.

Accordingly, here we followed Treisman’s (Treisman & Souther, 1985; Treisman & Gormican, 1988) visual search asymmetry strategy to determine the mental encoding of sex from body shape. Specifically, we used this approach to test whether the female body form is represented with reference to a default male body shape. If so, observers should more efficiently detect female body targets in arrays of male distractors than *vice versa*.

2.3 Experiment 1

2.3.1 Introduction

Participants either searched for a male body amongst female distractor bodies, or *vice versa*. Search efficiency was compared for the two target types by assessing the time required to detect a target over varying set sizes, and by sensitivity (d-prime) to detect targets of each type.

The bodies in this and all subsequent studies were rendered as silhouettes. Silhouettes have the advantage of capturing the shape of the body envelope without including confounding internal features such as clothing or texture. Silhouetted stimuli have been used successfully in previous studies of object (Wagenmans et al., 2008), face (Davidenko, 2007; Davidenko et al., 2012), and body (Downing et al., 2001; Downing et al., 2004; Stein et al., 2012; Gaetano et al., 2014; Palumbo et al., 2012; Thierry et al., 2006; Reeder et al., 2015; Reeder and Peelen, 2013) perception. In this and the following experiment, the body images were presented without the head and face. If the images were to include a face whose sex matched the body, then either or both cues could drive the hypothesized search asymmetry, rendering it difficult to interpret. If instead an ambiguous or “neutral” face were included, this would conflict with the sex signalled by the body, possibly interfering with typical sex judgment processes. Therefore, by presenting bodies only, we were able to exclude some other factors that would complicate interpretation of the results.

2.3.2 Method

2.3.2.1 Participants. Participants in this and the following experiments were students at Bangor University. They took part in return for course credit in a research methods module. No individual participant took part in more than one experiment. The procedures were approved by the Research Ethics Committee of Bangor University’s School of Psychology. Participants with mean response times or accuracy (averaged across conditions) of more than 2.5 SD below or above the group mean were considered outliers. Their data were excluded and new participants were tested to replace them to reach a sample size of N=32 (1 in Experiment 1; 2 in Experiment 2; 3 in Experiment 3; 3 in Experiment 4; 2 in Experiment 5; 2 in Experiment 6). This sample size was based on estimates from previous literature and was pre-registered along with

other aspects of the procedure of Experiment 1 here:

<http://aspredicted.org/blind.php?x=km9zi7>.

The final group of participants in Experiment 1 comprised 32 students (16 female) with a mean age of 24 years \pm 5.38 SD.

2.3.2.2 Apparatus and Stimuli. Images of leftward-facing, headless body profile silhouettes were generated using Poser (Curious Labs, Santa Cruz, CA). The images were black on a white background; each was 180 x 180 pixels, covering ~1 cm (width) x ~5 cm (height) on screen. The image set comprised 42 silhouettes (21 males, 21 females) that varied moderately in their overall body shape (**Figure 2.1a**). All images from this and the subsequent experiments are available here:

https://osf.io/cjvqd/?view_only=1dbde3b7a9364e52bf053d904eda76fd .

The experiments were administered using the Psychtoolbox (Brainard et al., 1997) package running in Matlab (MATLAB Release 2012b, The MathWorks, Inc., Natick, Massachusetts, US) on an Apple iMac computer. Viewing distance was approximately 60 cm from the screen but was not fixed.

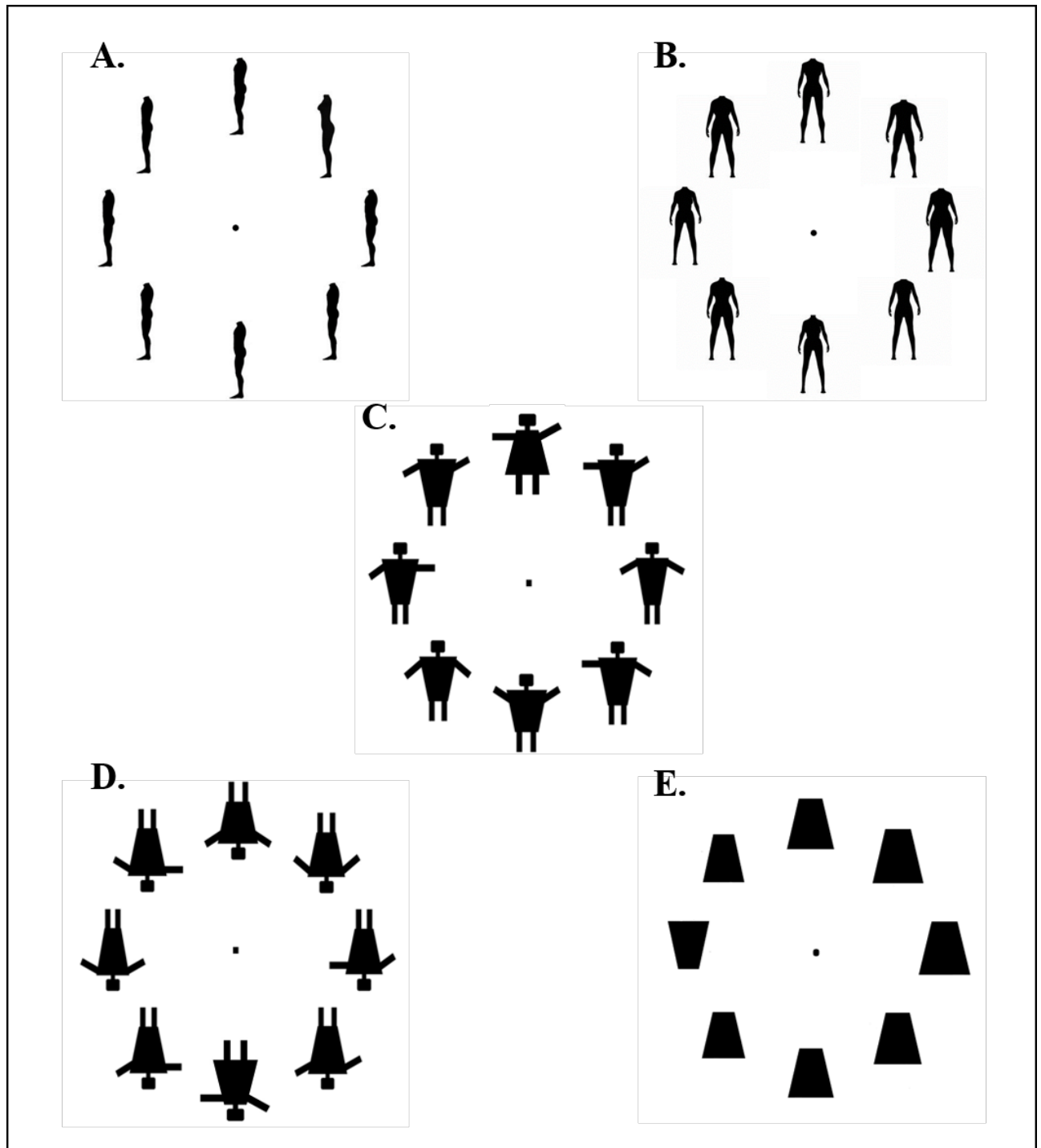


Figure 2.1. Example stimuli from Experiments 1-6 (A-E). Search set sizes included 1, 2, 4, or 8 items in Experiment 1-5; 4 or 8 in Experiment 6. Experiment 1 tested side views of body silhouettes. Experiment 2 tested frontal silhouettes. Experiment 3 tested curvature-less iconic body figures. Experiment 4 tested inverted iconic body figures. Experiment 5 tested displays consisting only of the geometric shapes (trapezia) that distinguished the sexes in Experiment 3 and 4. In Experiment 6, the same figures shown in panel C were tested, in an odd-one-out search task.

In the raw output of the Poser software, the male silhouettes occupied more surface area than the females. Visual size *per se* may be a veridical covariate of sex in viewed bodies (men, on average, being larger than women). However, it is also known from previous search asymmetry studies (Treisman and Gormican, 1988) that targets that comprise “more” visual content (e.g. longer lines, pairs of lines) are found more readily amongst distractors with “less” content (e.g. shorter lines, single lines) than *vice versa*. Hence we sought to control the stimulus size difference in the present study, so that if the predicted asymmetric pattern is observed, it can be attributed to the hypothesised differences between the body shapes taken by the two sexes, rather than size. Accordingly, the male silhouette images were rescaled, so that the proportion of black:white pixels in the items from the two categories did not differ on an independent samples t-test, $t(40) = 0.6$, $p = 0.55$.

2.3.2.3 Design and Procedure. Participants were instructed in different blocks either to search for a female body silhouette amongst male distractors, or a male body silhouette amongst female distractors. The design included four blocks, each comprising 128 trials; in two blocks the target was male, and in two female. The four blocks were presented in a counterbalanced order (MFFM or FMMF, equally across participants). Within blocks, trial orders were randomised such that each chunk of 16 trials consisted of two trials each from the crossing of target (present, absent) by set size (1, 2, 4, or 8). Participants were encouraged to take a short break between blocks.

Each trial started with a central fixation cross of random duration between 800 and 1200 ms. In each trial, 1, 2, 4 or 8 body silhouettes were presented for 3 seconds or until the participant responded. Each body silhouette could appear randomly in one of 8 possible locations that fell on a virtual circle (radius ~6 cm) around the fixation point. The distractors were randomly chosen from the image set and no body distractor could appear more than once in a given trial. The target, selected at random from the relevant item set, was present in 50% of the trials. Participants were instructed to “press J if a male [female] is present, press F if no male [female] is present”. Participants were encouraged to respond quickly without sacrificing accuracy.

2.3.2.4 Data Analysis. In all experiments we examined the efficiency of search considering two measures: search rates and sensitivity. Search rates were determined by estimating with a linear fit the slope relating search set size to response times (RT)

for present targets on accurate trials. Smaller values (flatter slopes) reflect more efficient search for the target. Sensitivity was assessed by computing d-prime measures (Macmillan and Creelman, 1990), collapsing over set sizes. Higher d-prime reflects better discriminability of a target present trial from a target absent trial. Extreme hit and false alarm proportions ($p = 0$ or $p = 1$) were adjusted according to Hautus (1995). Search slopes for male targets amongst female distractors were compared to slopes for female targets amongst male distractors with paired-samples two-tailed t-tests. Analogous contrasts were applied to the d-prime measures. As an exploratory analysis, we also determined a measure of response bias (β) from the sensitivity analysis. The value for β is defined as the ratio of the height of the signal plus noise distribution at the criterion to the height of the noise distribution at the criterion. As β gets larger the observer is said to be more conservative. As the criterion gets smaller β gets closer to 0 the observer is considered to be liberal. Criterion and d-prime were calculated in R (Version 3.6) using the package “Psycho” (Makowski, 2018) where the signal detection theory methods are calculated using the algorithms of Pallier (2002).

The mean slopes by target sex for each experiment are reported in **Figure 2.2**, the mean d-prime (and criterion) by target sex for each experiment in **Figure 2.3**, and plots of mean accurate RT against set size, as a function of target sex and target presence, are shown in **Figure 2.4**. We report mean accuracy by condition for each experiment in **Table 1**, but do not analyse it further given that it was generally near ceiling. However, in no study was there evidence indicating a speed-accuracy trade-off.

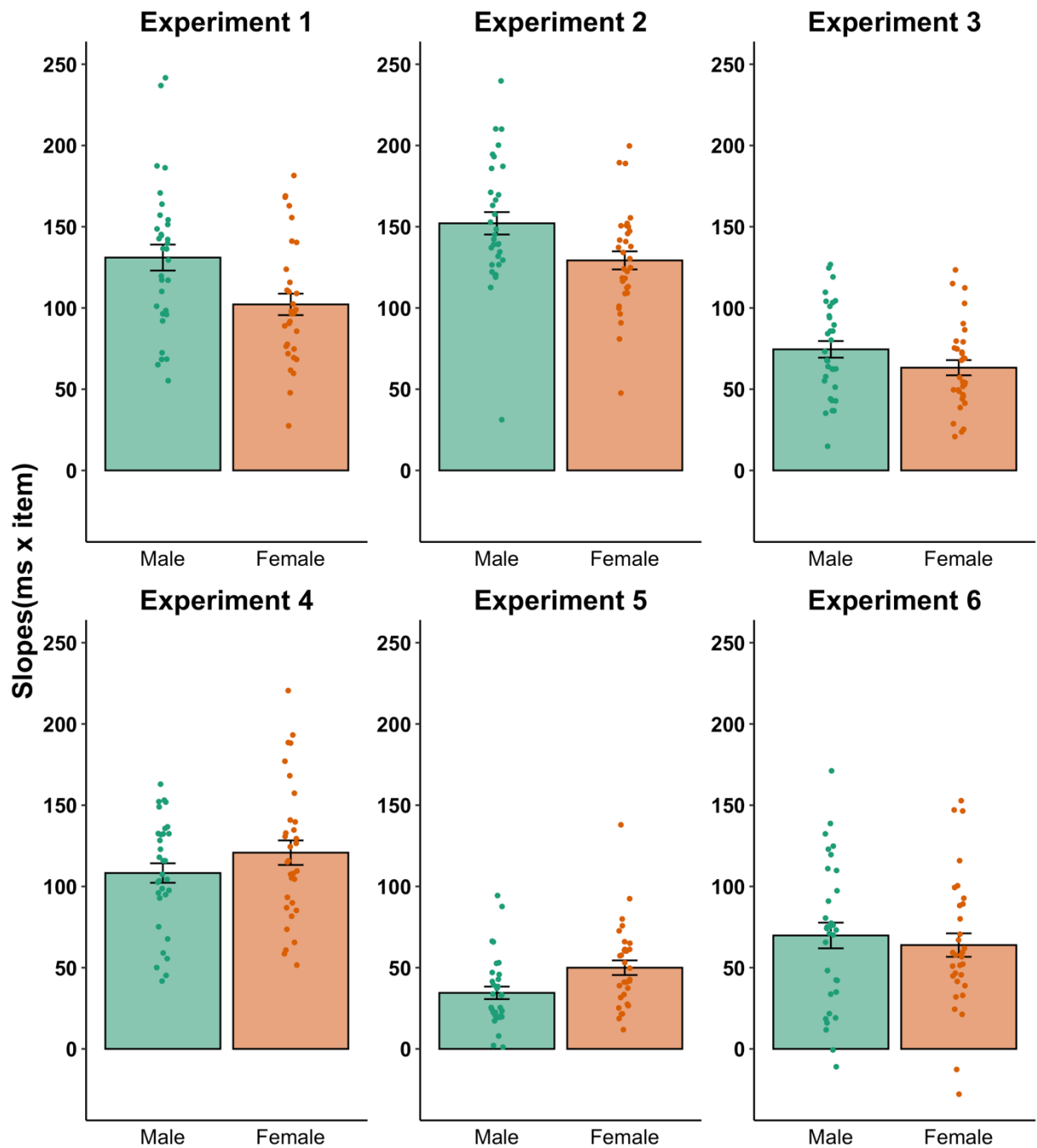


Figure 2.2. Summary of results of Experiments 1-6. Efficiency of visual search for male targets amongst female distractors, and vice versa, is shown in terms of ms/item derived by fitting a linear function to search times on accurate, target-present trials. Error bars represent ± 1 SEM, * $p < 0.05$; ** $p < 0.01$. Search for females is more efficient across side and frontal views (Experiments 1 and 2) and also in iconic figures that lack curvature (Experiment 3). In contrast, search for the inverted iconic figures (Experiment 4) does not show a significant difference in search efficiency. Search for the individual shape elements that distinguished male from female

in Experiment 3 showed the opposite pattern (Experiment 5). No significant difference on search slopes was found in experiment 6.

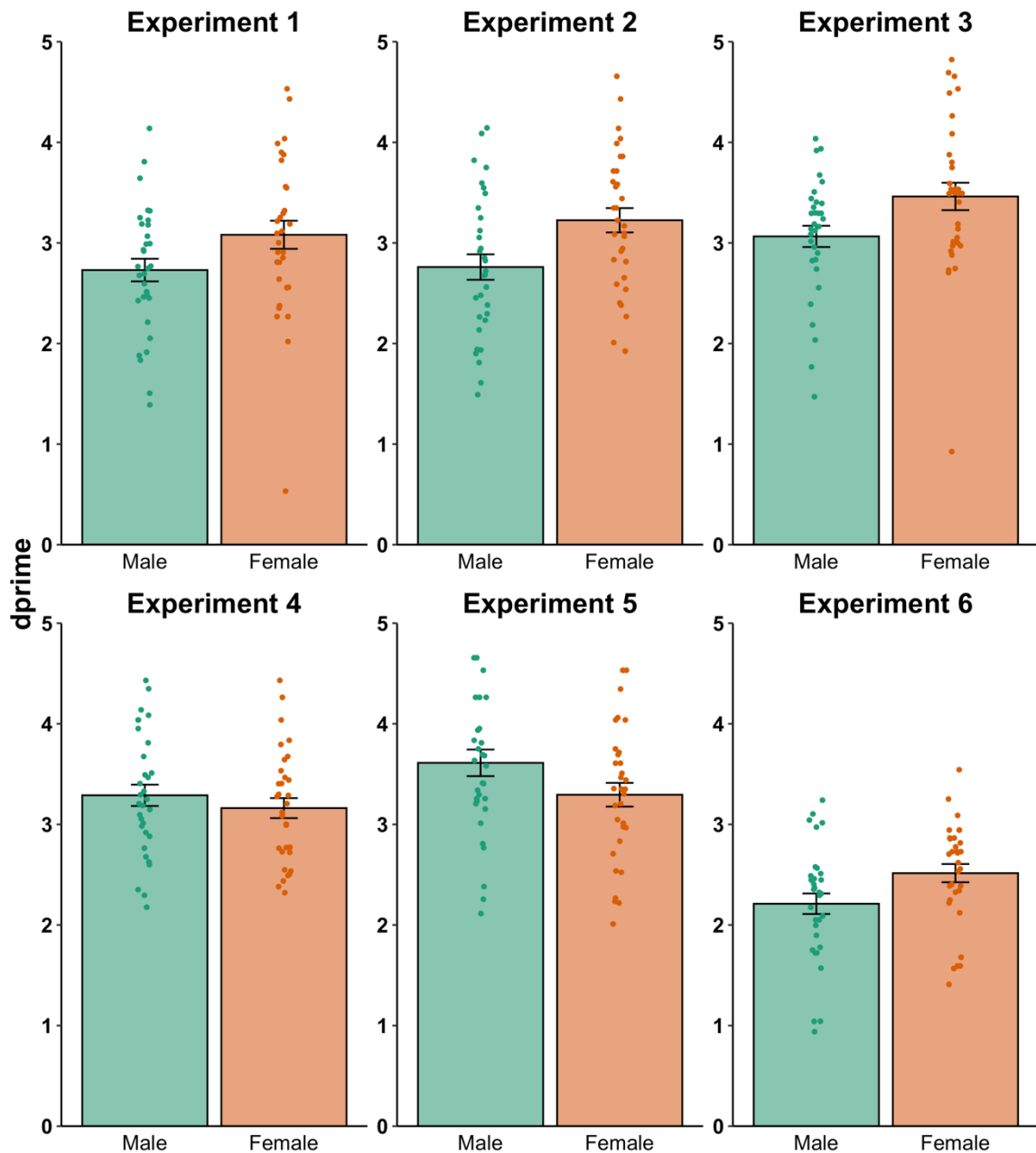


Figure 2.3. Summary of results of Experiments 1-6. Accuracy of visual search for male targets amongst female distractors, and vice versa, is shown in terms of d-prime. Error bars represent ± 1 SEM, * $p < 0.05$; ** $p < 0.01$. Sensitivity for females is more efficient across side and frontal views (Experiments 1 and 2) and also in iconic figures that lack curvature (Experiment 3), also when participants were tasked to find the odd-one out for sex (Experiment 6). In contrast, search for the inverted iconic figures (Experiment 4) does not show a significant difference in

search efficiency. Search for the individual shape elements that distinguished male from female in Experiment 3 showed the opposite pattern (Experiment 5).

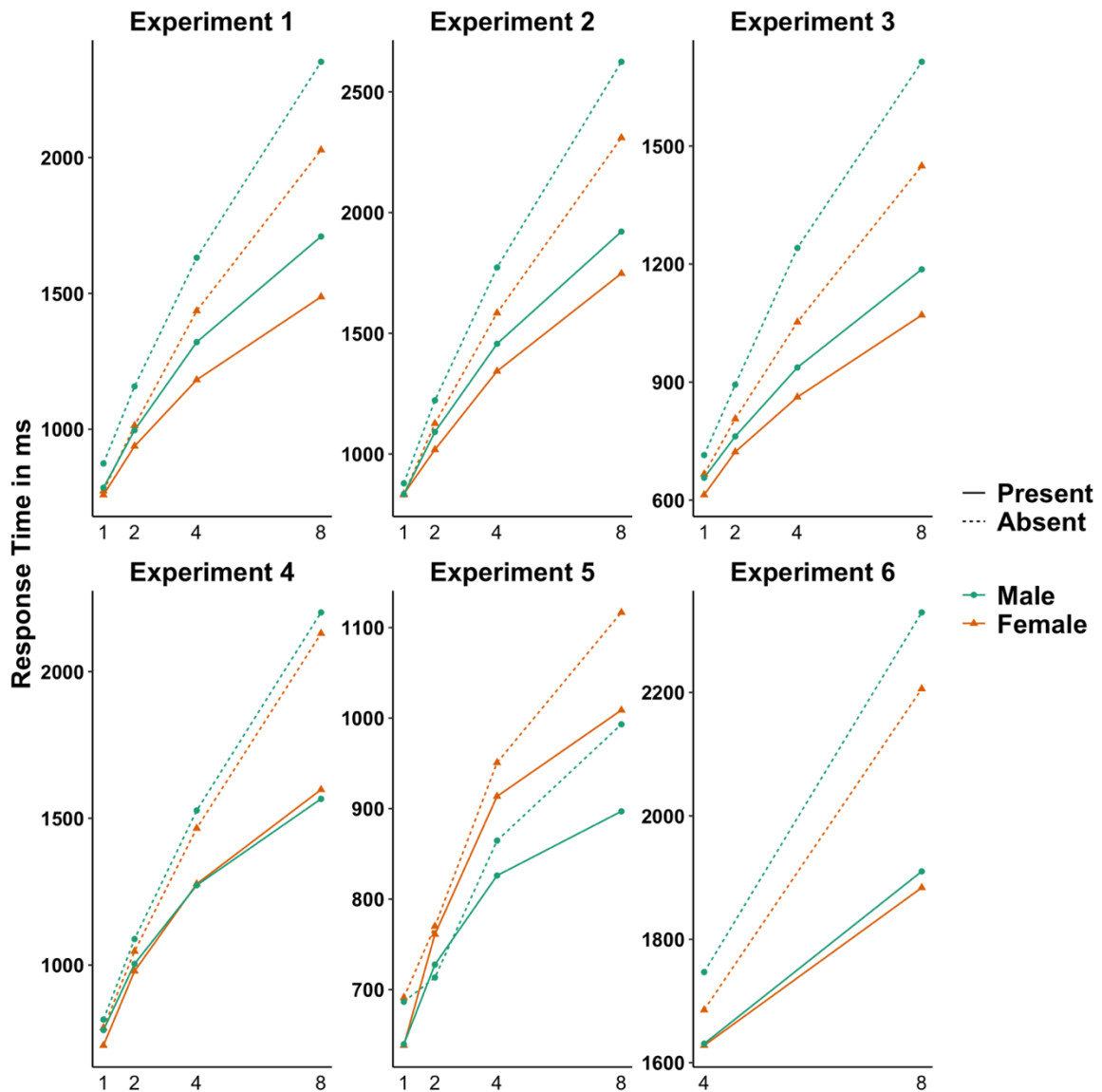


Figure 2.4. Mean response times from target-present (solid line) and target-absent (dashed line) trials, in which either male (square marker, green line) or female (circle marker, orange line) body shapes were the search targets for Experiment 1-5 (A-E).

2.3.3 Results

A paired samples *t*-test on search slopes for target present trials showed a significant difference between searching for male vs female bodies, $t(31) = 4.98$, $p < 0.001$, $d = 0.88$, 95% CI = [16.7, 45.6]. (Here and in the following experiments, confidence intervals are calculated over the difference between male and female target conditions). Search for female body silhouettes among male distractors ($M = 100$ ms/item; $SE = 6.84$) was more efficient than for males among females ($M = 128$ ms/item; $SE = 8.19$). A paired-samples *t*-test on *d*-prime showed significantly higher sensitivity when searching for a female target ($M = 3.08$, $SE = 0.14$) vs a male target ($M = 2.76$, $SE = 0.11$), $t(31) = 2.71$, $p = 0.01$, $d = 0.48$, 95% CI = [0.09, 0.61]. A more conservative criterion was observed when searching for female ($M = 7.15$, $SE = 0.43$) vs male target ($M = 2.17$, $SE = 0.43$), $t(31) = 3.14$, $p = 0.004$, $d = 0.55$, 95% CI = [1.74, 8.22].

2.3.4 Discussion

As predicted by the hypothesis that the shape of the female body is represented as a departure from a reference male body shape, females were detected more rapidly and more sensitively than were males, amongst distractors of the opposite sex. Furthermore, the high mean level of performance on the task, as indicated by the *d*-prime results, confirms that participants were readily able to categorise the silhouettes in this study as female or male.

2.4 Experiment 2

2.4.1 Introduction

In Experiment 2, we assessed the generality of the previous finding by repeating the study with body silhouettes shown from a frontal view. Frontal and lateral views differ in respect to the aspects of body shape that they reveal or obscure. To the extent that the same search asymmetry is found again, this indicates that it is less likely to be driven by any specific idiosyncratic property of the images in either stimulus set.

2.4.2 Methods

Except as noted, the methods of Experiment 2 were identical to Experiment 1. The final sample of participants included 32 students (16 female; mean age 24 years \pm 5.84).

Images of frontal body silhouettes were generated using the open source software MakeHuman (Version 1.0.1, www.makehuman.org). The depicted bodies held a neutral standing pose with arms open, and were cropped to exclude the head (Figure 2.1B). As in Experiment 1, by cropping the head from the images, we ensured that any asymmetry in search efficiency found in this study would be attributable to body shape representations specifically, and not to potentially confounding aspects of the head or face. The pool of images (link: https://osf.io/cjvqd/?view_only=1dbde3b7a9364e52bf053d904eda76fd) comprised 32 body silhouettes (16 males, 16 females) that differed in body shape. These were obtained generating random MakeHuman characters (with variable weight and body proportions across sexes) and applying the phong shader (with emissive set to black) embed in the material editor of the software. Surface area of the silhouettes was controlled as in Experiment 1, and did not differ between sexes, $t(30) = 1.25$, $p = 0.22$.

2.4.3 Results

A paired samples t-test on search slopes for target present trials showed a significant difference between searching for male vs female bodies, $t(31) = 4.14$, $p < 0.001$, $d = 0.73$, 95% CI = [11.3, 33.3]. Search for female body silhouettes amongst male distractors ($M = 128$ ms/item; $SE = 5.74$) was more efficient than for males amongst females ($M = 150$ ms/item; $SE = 6.88$). A paired-samples t-test on d-prime showed significantly higher sensitivity when searching for a female target ($M = 3.23$, $SE = 0.12$) vs a male target ($M = 2.76$, $SE = 0.13$), $t(31) = 4.70$, $p < 0.001$, $d = 0.83$, 95% CI = [0.26, 0.67]. A more conservative criterion was observed when searching for female ($M = 6.87$, $SE = 1.28$) vs male target ($M = 3.33$, $SE = 0.67$), $t(31) = 2.73$, $p = 0.01$, $d = 0.48$, 95% CI = [0.90, 6.20].

As a *post hoc* test of whether the participants' gender influenced asymmetries in search efficiency, we combined the data on slopes from present targets from Experiments 1 and 2 in a mixed design ANOVA with participant gender and target sex as factors. Search was more efficient for female than for male body targets, $F(1,62) =$

41.89, $p < 0.001$, $\eta p^2 = 0.40$, and female participants were more efficient in the search task ($M = 118$ ms/item, $SE = 6.58$) than were male participants ($M = 136$ ms/item, $SE = 6.55$), $F(1,62) = 4.09$, $p = 0.047$, $\eta p^2 = 0.06$. However, the interaction of these two variables was not significant, $F(1,62) = 0.61$, $p = 0.44$, $\eta p^2 = 0.001$. The same analysis on sensitivity showed a higher discriminability for female than male body targets $F(1,62) = 25.19$, $p < 0.001$, $\eta p^2 = 0.29$ but not a main effect of sex $F(1,62) = 0.92$, $p = 0.34$, $\eta p^2 = 0.01$ or a target sex x participant's sex interaction $F(1,62) = 0.85$, $p = 0.36$, $\eta p^2 = 0.01$.

2.4.4 Discussion

Experiment 2 revealed a search asymmetry for body sex as did Experiment 1, demonstrating that this result generalises across viewpoint. This pattern did not differ reliably for female and male participants, so in subsequent experiments we made no further effort to balance for gender when recruiting participants.

In both experiments, the properties that distinguish the two sexes are, at least in part, related to curvature, given that most other visual features are absent from silhouettes. Studies of search for simple visual elements have found that search for a curve among straight lines is more efficient than the converse (Treisman and Gormican, 1988). This finding was interpreted as indicating that curvature is encoded as a deviation relative to a linear standard. Moreover, there is previous evidence for search asymmetries favouring concavities over convexities (Hulleman et al., 2000), and favouring shapes that contain curvature discontinuities (a local change in the rate of curvature) relative to those that do not (Kristjánsson and Tse, 2001). These findings are particularly relevant for Experiments 1 and 2, in that silhouettes of female bodies may have more (or more pronounced) concavities or curvature discontinuities than do silhouettes of male bodies, which could in turn explain at least part of those findings. These considerations motivated the following experiments.

2.5 Experiment 3

2.5.1 Introduction

In Experiment 3 we examined search amongst iconic body images consisting only of linear elements. In these figures, the bodies did not contain continuous curves, thereby eliminating potential confounding differences in curvature between the two

conditions. Furthermore, the stimuli were constructed such that identical geometric elements were arranged in different configurations in order to produce a percept of sex from body shape. In this way other potential confounds of visual properties between the two conditions were eliminated. For example, with this manipulation we can be sure that the items from the two sexes are equivalent with respect to the homogeneity of the exemplars – that is, the variance in shape across the exemplars within each class. Although any such differences in Experiments 1 or 2 were likely to be subtle if present at all, homogeneity may influence visual search efficiency: a given target type should be easier to find to the extent that the distractors are homogenous, other factors being equal (Duncan and Humphreys, 1989).

The additional experimental control over confounding variables that is offered by testing iconic bodies comes with a tradeoff against ecological validity relative to the more naturalistic stimuli of Experiments 1 and 2. While (as will be seen below) the iconic figures were readily accepted and identified by participants as human figures with a defined sex, they are not naturalistic in appearance. However, on this point it is worth noting the extensive previous literature on other aspects of social vision, particularly work on faces, in which averaged, filtered, cropped, edited, schematic, cartoon, or caricatured images have been used to great effect in order to both control stimulus confounds and to test key hypotheses.

2.5.2 Methods

Except as noted, the methods of Experiment 3 were identical to Experiments 1 and 2. The final sample included 32 students (25 female; mean age 19 years \pm 1.93 SD).

Icon images of human male and female bodies were prepared using PowerPoint (Microsoft Office, 2016; Figure 2.1C). To begin, we created 8 upward-pointing trapezia (narrower at the top) and 8 downward-pointing trapezia (narrower at the bottom) of varying proportions. Upward-pointing trapezia were treated as the torso for female figures, and downward-pointing trapezia as torsos for male figures. These shapes were then flipped around the horizontal axis, in order to create 8 more torsos of each sex, resulting in a final set of 32 images (16 males, 16 females). These torsos were elaborated with rectangles representing legs, arms, neck and a head. Pose was varied such that the arms variously pointed up, down, or to the side (matched across the two

sexes). The surface area of the figures did not differ between sexes, $t(30) = 0.77$, $p = 0.44$.

2.5.3 Results

A paired samples t-test on search slopes for target present trials showed a significant difference between searching for male vs female icon bodies, $t(31) = 2.24$, $p = 0.03$, $d = 0.40$, 95% CI = [1.02, 21.6]. Search for female body icons amongst male distractors ($M = 63$ ms/item; $SE = 4.68$) was more efficient than for males amongst females ($M = 74$ ms/item; $SE = 5.12$). A paired-samples t-test on d-prime showed significantly higher sensitivity when searching for a female target ($M = 3.46$, $SE = 0.14$) vs a male target ($M = 3.06$, $SE = 0.11$), $t(31) = 4.21$, $p < 0.001$, $d = 0.74$, 95% CI = [0.20, 0.59]. No difference in the decision bias between the two tasks was found, $t(31) = 0.80$, $p = 0.43$, $d = 0.14$, 95% CI = [-0.58, 1.33].

The rapid search rates and high d-prime values observed here confirm that our participants accepted these icon silhouettes as representations of male and female bodies, and were able to categorise them efficiently.

2.5.4 Discussion

The effect of sex on efficiency of search amongst body images does not depend entirely on the presence of naturalistic curvature in the image. Even in curvature-less icon stimuli, search for female figures amongst males was more efficient than the converse. The following two studies pursued this finding, with aims to further rule out possible confounding factors, and to examine the relative contributions of local body elements vs holistic shape to sex judgments. Specifically, we tested whether the search asymmetry persists for inverted figures (Experiment 4) and whether it can be explained entirely by the orientation of the central “torso” element of the icon figures (Experiment 5).

2.6 Experiment 4

2.6.1 Introduction

Inverting a face (Rossion and Gauthier, 2002; Yin, 1969; Yovel and Kanwisher, 2005) or a body (Matsuno and Fujita, 2018; Minnebusch et al., 2009; Minnebusch et al., 2010; Reed et al., 2003; Reed et al., 2006; Stein et al., 2012; Sumi, 1984; Yovel et al., 2010) in the image plane can result in disproportionate impairment of perception of these stimuli, relative to other kinds of objects. For bodies, it is proposed by several authors that inversion interrupts normal whole-body perceptual processes, so that perceivers must instead rely more on a local, part-based analysis of body images to perform visual tasks (Robbins and Coltheart, 2012; Stekelenburg and Gelder, 2004; Urgesi et al., 2007). Therefore, inversion provides an opportunity to better understand the body representations that underlie the search asymmetry observed in the preceding studies (and particularly in Experiment 3). If the search advantage for female bodies persists for inverted icon figures, then this would suggest that the effect is driven, at least in part, by local or part-based aspects of the body shape. In contrast, if inversion reduces or eliminates the search asymmetry, this would suggest instead a process operating at the whole-figure level.

Finally, if the same search asymmetry is not found for inverted icon figures, this would provide compelling evidence that the asymmetry found for the same icons presented upright cannot be attributed to a confounding (orientation-invariant) low-level property of the stimuli.

2.6.2 Methods

Except as noted, the methods of Experiment 4 were identical to the preceding experiments. The final sample included 32 students (22 female; mean age 19.27 years \pm 1.85). The stimuli for Experiment 4 consisted of the icons from Experiment 3 after rotating them by 180° (**Figure 2.1d**).

We sought to ensure that participants were aware of the status of the inverted icons as depictions of human bodies. Were they not aware, then a failure to detect a search asymmetry in this study might have been attributable to a difference between experiments in participants' understanding of the images. Accordingly, participants were first briefly shown examples of the upright body icons on paper, and were told that they had been used in a previous study about perception of sex from the body. They were then told that they were taking part in a study in which those body pictures would

appear upside down; the paper was then rotated to show how these inverted body figures would look in the task. Participants were then instructed as in the previous experiments to search in different blocks for either a female body icon among male distractors, or a male body icon among female distractors.

2.6.3 Results

A paired-samples t-test on search slopes for target present trials with accurate responses (female: 0.88, SE = 0.01; male: 0.92, SE = 0.01) revealed no significant search asymmetry between inverted female ($M = 116$ ms/item; SE = 7.68) and inverted male icons ($M = 108$ ms/item; SE = 6.03), $t(31) = 1.45$, $p = 0.15$, $d = 0.25$, 95% CI = [-26.01, 4.32]. A paired-samples t-test on d-prime showed no difference in sensitivity when searching for a female target ($M = 3.16$, SE = 0.10) vs a male target ($M = 3.23$, SE = 0.11), $t(31) = 1.33$, $p = 0.19$, $d = 0.24$, 95% CI = [-0.32, 0.07]. No difference in the decision bias between the two tasks was found, $t(31) = 1.55$, $p = 0.13$, $d = 0.27$, 95% CI = [-0.41, 3.04].

A direct comparison revealed that this pattern was different from the search asymmetry found in Experiment 3: the interaction of Orientation (between participants) x Target Sex (within participants) was significant on search rates, $F(1, 62) = 6.08$, $p = 0.02$, $\eta^2 = 0.09$, and on d-prime $F(1, 62) = 15.32$, $p < 0.001$, $\eta^2 = 0.20$. Further, as expected, search in target present trials was more efficient overall for upright icons (Experiment 3; $M = 68$ ms/item; SE = 4.21) than for inverted icons (Experiment 4; $M = 110$ ms/item, SE = 5.81), $F(1, 62) = 34.60$, $p < 0.001$, $\eta^2 = 0.35$. No overall difference in sensitivity was found between upright and inverted icons ($p > 0.05$).

2.6.4 Discussion

Inverting the body icons increased task difficulty and removed the advantage for detecting female amongst male bodies seen in the preceding experiments. This finding suggests that, in line with the previous findings noted in the Introduction to this experiment, normal body perception processes were interrupted here by inversion. More specifically, following the assumption that inversion interferes with normal processing of the body image as a whole entity, the present finding suggests that the search asymmetry favouring female bodies lies in emergent characteristics of the whole

figure. This is as opposed to a basis in perception of specific local elements, whose processing we assume to be relatively unimpaired by inversion.

There is previous evidence that body inversion effects are only present, or are stronger, when the head or face is also included in the stimulus (Brandman and Yovel, 2012; Minnebusch et al., 2009; Yovel et al., 2010) although the underlying reasons for this finding are still being revealed (Arizpe et al., 2017; Robbins and Coltheart, 2012). Because the icons used in the present study included a simple “head”, we assume that at least in this respect they are suitable to generate inversion effects in line with those previous findings.

Finally, a benefit of directly comparing search for upright and inverted versions of the same stimuli is that they are perfectly matched for visual properties at the pixel-by-pixel level. Therefore, this manipulation rules out any account of the search asymmetry that would seek to explain it entirely on the basis of (orientation-invariant) lower-level stimulus properties that may be confounded between the male and female figures tested here.

2.7 Experiment 5

2.7.1 Introduction

The male and female icons tested in Experiments 3 and 4 differed only in the orientation of the trapezium that formed their “torsos”. Although the results of Experiment 4 would appear to rule out this possibility, it may be that this variation in itself, rather than perceived sex, influenced search efficiency. To test this directly, we repeated the experiment with figures that contained only the “torso” element (and with no reference to sex in participants’ instructions). If a similar asymmetry were observed, such that the upward-pointing trapezia (which were the basis for the female icons) were found more rapidly amongst downwards-pointing trapezia than vice versa, it would suggest that geometric shape rather than perceived sex best explains the results of Experiment 3. However, a previous finding suggests the opposite prediction: visual search tasks with triangles show an asymmetry favouring downward-pointing triangles amongst upward-pointing triangles (Larson et al., 2007).

2.7.2 Methods

Except as noted here, the methods of Experiment 5 were identical to the preceding experiments. The final sample included 32 students (26 female; mean age $20.71 \text{ years} \pm 3.38$). The stimuli for Experiment 5 consisted of the icons from Experiments 3 with all other elements aside from the central “torso” removed (Figure 2.1e). An independent samples t-test between upright and inverted trapezia pictures confirmed that they did not differ in the proportion of black pixels, $t(30) = 0.59$, $p = 0.56$. (That they were not strictly identical was due to small variances introduced in resizing and jpeg-compressing the images from the original PowerPoint figures).

At the start of each block, an example of a target shape and a distractor shape were shown to participants in the instructions that informed them how to respond. Images of targets and distractors were used rather than verbal labels, to avoid using descriptions such as “upward-pointing” that could bias participants’ framing of the stimuli and hence search efficiency. No mention of body perception was made in the study recruitment materials nor in the task instructions.

2.7.3 Results

A paired samples t-test on search slopes for target present trials showed a significant difference between searching for upward- vs downward-pointing trapezia, $t(31) = 3.63$, $p = 0.001$, $d = 0.64$, 95% CI = [6.74, 24.0]. Search for downward-pointing shapes ($M = 34 \text{ ms/item}$; $SE = 3.81$) was more efficient than for upward-pointing shapes ($M = 49 \text{ ms/item}$; $SE = 4.36$). A paired-samples t-test on d-prime showed a difference in sensitivity when searching for a downward-pointing shape ($M = 3.61$, $SE = 0.13$) vs a upward-pointing shape ($M = 3.29$, $SE = 0.12$), $t(31) = 3.44$, $p = 0.002$, $d = 0.61$, 95% CI = [0.13, 0.50]. No difference in the decision bias between the two tasks was found, $t(31) = 0.43$, $p = 0.67$, $d = 0.08$, 95% CI = [-1.21, 0.79].

2.7.4 Discussion

The search asymmetry found in Experiment 3 cannot be explained by the orientation of the “torso” element of the body icons, which was the only difference between the bodies of the two sexes in that study. We conclude instead that the asymmetric pattern of search efficiency for iconic figures must be due to distinctions in the representations of male and female bodies that those icons engage.

In fact, the asymmetry for shapes in isolation was in the opposite direction, such that the upward-pointing trapezia that formed the female torsos in Experiment 3 were detected amongst distractors less efficiently than the downward-pointing shapes. This is consistent with a previously reported attentional advantage for simple shapes containing a “V” or downward-pointing triangle (Larson et al., 2007). In other words, comparing across Experiments 3 and 5 we can see that the influence of the orientation of these central geometric elements is strongly driven by the context in which they appear. Viewing them in the context of a head and limbs leads to them being interpreted as part of a body, with the effect of reversing their relative influence on search efficiency.

2.8 Experiment 6

2.8.1 Introduction

Does the search asymmetry observed here depend on participants searching for a target that is defined explicitly, in advance, by sex? In the preceding experiments, the search target type was fixed for blocks of trials. This design permits both top-down guidance (participants are able to maintain a perceptual template of the target over a series of trials) and priming effects (targets and distractor types repeat over a series of trials), either or both of which may influence search efficiency. If the search asymmetry favouring female targets is entirely dependent on these factors, then this pattern should not be found in a task in which the status of each sex as a target or distractor varies unpredictably. If instead a similar asymmetry is observed again, this would suggest that the underlying mechanisms relate more closely to immediate stimulus-driven influences of body shape.

Accordingly, in Experiment 6 participants performed an “odd-one out” task on the icon figures of Experiment 3. The search arrays in this task were either homogenous with respect to sex (all male or all female; target absent) or else had a single deviating item (one male amongst females or vice versa; target present). These four trial types were randomised within blocks, so that from trial to trial participants could not anticipate whether male or female items would serve as target or distractor. Therefore, the only viable search strategy was to examine the items for a figure that deviated from the others by its sex.

2.8.2 Methods

Except as noted, the methods of Experiment 6 were identical to Experiment 3. The final sample included 32 students (25 female; mean age 21.03 years \pm 3.55). The stimuli of Experiment 3 were used. Participants were instructed to indicate on each trial whether or not there was an odd item, with respect to sex, in the search array. Because “oddity” is not defined for set sizes of 1 or 2 we only tested the larger set sizes of the previous experiments (4 or 8). The design included 256 trials; trial orders were randomised such that each chunk of 16 trials consisted of four trials from the crossing of target (present, absent) by set size.

2.8.3 Results

A paired-samples t-test on search slopes for target present trials with accurate responses (female: 0.86, SE = 0.06; male: 0.84, SE = 0.06) showed no significant asymmetry in search efficiency when searching for an odd male target ($M = 70$ ms/item, SE = 7.90) compared to an odd female target ($M = 64$ ms/item, SE = 7.22), $t(31) = 0.64$, $p = 0.53$, $d = 0.11$, 95% CI = [-24.80, 13.00]. A paired-samples t-test on d-prime showed significantly higher sensitivity when searching for an odd female target ($M = 2.52$, SE = 0.09) vs an odd male target ($M = 2.21$, SE = 0.10), $t(31) = 4.34$, $p < 0.001$, $d = 0.77$, 95% CI = [0.16, 0.45]. No difference in the decision bias was found between the two tasks $t(31) = 1.59$, $p = 0.12$, $d = 0.28$, 95% CI = [-0.09, 0.69].

2.8.4 Discussion

The search asymmetry found in the previous studies cannot be entirely explained by top-down effects, or priming effects, arising from the specification of a fixed target sex in each block. In this “oddity” task, female targets were still more discriminable among male targets than vice-versa, even though target and distractor roles varied unpredictably from trial to trial. These findings support the hypothesis that the search asymmetry we find reflects the perceptual organization of body sex representation and is not entirely a consequence of top-down guidance processes (Wang et al., 2005; Friedman et al., 1995) nor of trial-by-trial priming effects. That said, we note that unlike in Experiments 1-3, the search asymmetry was revealed primarily in the d-prime analysis of sensitivity rather than in efficiency as measured by RT slopes. This may in part be accounted for by task difficulty – judging from the d-prime measures and the RT curves, the oddity task was more difficult than all of the preceding experiments. In a

more difficult search task, for example, observers may spend a greater proportion of the search time re-examining individual items, hence diluting the impact of the stimuli on measures of slope. Further, the set-size in this experiment consisted of 4 or 8 elements only. A smaller range of set sizes might hinder the sensitivity of this variable in capturing search efficiency.

Table 2.1. Descriptive statistics for Accuracy in each experiment and target sex.

Experiment	Target	<i>M</i>		
		<i>M</i>	95% CI [LL, UL]	<i>SD</i>
Experiment 1	Female	0.86	[0.83, 0.90]	0.10
	Male	0.87	[0.84, 0.90]	0.08
Experiment 2	Female	0.87	[0.84, 0.90]	0.08
	Male	0.87	[0.84, 0.89]	0.08
Experiment 3	Female	0.93	[0.91, 0.96]	0.07
	Male	0.91	[0.90, 0.93]	0.05
Experiment 4	Female	0.88	[0.85, 0.91]	0.08
	Male	0.92	[0.90, 0.94]	0.06
Experiment 5	Female	0.92	[0.90, 0.94]	0.06
	Male	0.94	[0.91, 0.96]	0.06
Experiment 6	Female	0.86	[0.83, 0.89]	0.08
	Male	0.85	[0.82, 0.87]	0.08

2.9 General Discussion

Visual search for female bodies among male distractors is more efficient than the converse. This finding generalizes across views, and does not depend on the presence of curvature discontinuities in the images, on the presence of the head in the stimuli, or

on the gender of the participants. When examining visual perception of complex social stimuli, it is not always possible to completely rule out the contribution of confounded image properties. However, this concern is mitigated here by several factors: the use of silhouettes, which exclude many co-varying aspects of colour, texture, and shading; matching silhouettes for surface area, which could in itself influence search efficiency; testing silhouettes from frontal and side views, as well as in an iconic format; finding a different pattern of search efficiency for otherwise identical upright and inverted icon figures; and a control experiment ruling out a geometric account of the results from those icon figures.

The asymmetric search efficiency that we found is at least superficially similar to that seen in search for more basic visual elements, such as for Qs among Os or ellipses among circles. In some cases, such as Qs among Os, the asymmetry is pronounced enough that in the more efficient condition, search times are approximately set-size invariant (“pop-out”). In other cases, such as the present study, search rates are asymmetric, but neither target type pops out. This pattern is to be expected for bodies if we assume that the representations of female and male bodies share several properties in common, rendering them relatively similar and hence making selection relatively difficult (Duncan & Humphreys, 1989).

The search asymmetries strategy developed by Treisman (Treisman and Souther, 1985; Treisman and Gormican, 1988) and pursued since by other authors (Wolfe, 2001) has met with some criticism. In particular, Rosenholtz (Rosenholtz, 2001) argues that in some previous applications of the logic, particularly investigations of colour or motion analysis, what appears to be an asymmetric result in a symmetric design may actually relate to an asymmetric design. That is, for some designs, an apparently symmetric reversal of target and distractor roles may not be so, depending (for example) on the within-class heterogeneity of distractor items, or on the relationship of oriented items to the static reference frame of the surrounding environment. Such analyses depend on having clearly identifiable linear dimensions (e.g. location and velocity, or colour space) on which to define the search stimuli. At present, our understanding of the “space” of body representations is still in development (Hu et al., 2018), and indeed one aim of this work is to provide constraints on such models. More generally, we argue that the present studies meet the criteria for genuinely symmetric design as laid out by Rosenholtz (Rosenholtz, 2001).

How can we describe the properties of male and female body shapes that account for asymmetric search performance? Note that to do so is not trivial even for apparently more elementary cases. For example, are Q targets so readily found among O distractors because of the additional line segment itself, or the junction formed where it joins the curve of the letter, or the interruption of that curve, or the introduction of higher spatial frequencies, or all of these factors in combination? The question is still more complex for bodies, and the answer likely does not correspond to a single concrete visual characteristic being present or absent in the stimulus, given that we found the same pattern over different image formats (and not in inverted images). For example, in frontal views, one candidate property is waist-to-hip ratio (Singh, 1993), but this feature is obscured in silhouetted lateral views of the body, which also generated the search asymmetry.

Indeed, our finding with icons shows that the search asymmetry can be generated by implied features of the two sexes, because real bodies do not take these specific visual shapes. Similarly, previous studies show that search efficiency can be strongly influenced by implied features and by emergent properties. This includes evidence on search for shapes defined by illusory edges (Li et al., 2008); on the effects of congruence between individual visual elements with an implied three-dimensional arrangement of objects (Enns and Rensink, 1990); and on the (lack of) interference from visual surfaces that can be construed as cast shadows (Rensink and Cavanagh, 2004). In other words, search performance and search asymmetries are driven not only by the physical characteristics of the stimulus but also by implied real-world meaning.

We also know that brain regions and electrophysiological signatures that are closely tied to body perception (Minnebusch and Daum, 2009; Peelen and Downing, 2007; Sadeh et al., 2011) can be strongly and selectively engaged by minimal, iconic, or otherwise less-than-realistic body depictions. These include line drawings, silhouettes, stick figures, and point-light animations (Downing et al., 2001; Grossman et al., 2000; Kana and Travers, 2012; Peelen and Downing, 2005; Peelen et al., 2006; Schupp and Renner, 2011; Thierry et al., 2006; Uher et al., 2005). Therefore, there is neural evidence for the ready and rapid engagement of body representations by minimalistic stimuli similar to the silhouettes and icons tested here.

Drawing these considerations together, we argue that the present findings are not best explained by the presence or absence of a specific isolable visual element. Instead, they reflect the influence of a holistic, emergent representation of body shape that distinguishes female body form from a male “default”. That this representation generalises over views is consistent with encoding at a relatively high level of abstraction, in line with the properties of more anterior body-selective regions in the primate (Kumar et al., 2017) and human (Taylor et al., 2010) brain. Likewise, that this representation generalises over images of varying realism indicates that it captures abstract properties of the body that relate as much to the subjective percept as to the physical stimulus itself.

The latter claim suggests that the asymmetric pattern found here may depend in part on deliberate attention to sex as a category. Yet social-psychological perspectives on person perception have conceived categorisation by sex as automatic or obligatory (Stangor et al., 1992). If so, a search asymmetry favouring female over male bodies would also be found in tasks in which sex varies incidentally but is not the target-defining property. Conversely, adopting a goal to select targets on the basis of sex is not sufficient in its own right to generate the asymmetric pattern, as this was disrupted by inversion in Experiment 4. Finally, another way to disentangle the effects of top-down influences such as task set and subjective percept, from bottom-up stimulus factors, would be to ask whether sex from body shape influences detection efficiency before the stimulus reaches awareness, as tested in inattentional blindness and continuous flash suppression tasks (Downing et al., 2004; Stein et al., 2012).

One potentially relevant factor in shaping mental representations of body form is differential expertise, which may in turn shape visual search efficiency, as seen for faces and other objects (Golan et al., 2014; Hershler and Hochstein, 2009; McGugin et al., 2011; Reeder et al., 2016; Tong and Nakayama, 1999). For example, infants often show a preference for, or greater fluency with, female relative to male faces (Ramsey-Rennels and Langlois, 2006) and bodies (Alexander et al., 2016). One prominent view is that these findings relate to more extensive experience with female relative to male caregivers, which is typical in many societies (Ramsey, Langlois, & Marti, 2005; Rennels et al., 2017). For example, in one influential study infants aged 3 to 4 months looked longer at female faces when they were presented paired together with male faces (Quinn et al., 2002). Importantly, while this preference was not explained by

differences in cues related to hair or to image contrast, it did depend on the participants having females as primary caregivers, and was reversed in a sample of infants raised by male caregivers, in support of an expertise account.

However, to link such observations to a plausible account of the present findings in young adults would require additional assumptions that we have not tested. And, importantly, there is evidence that the influence of experience with caregivers on face perception wanes during puberty (Picci and Scherf, 2016). Further, attributing a search asymmetry to differential expertise does not in itself offer a mechanistic basis by which behavioural effects might be generated. But one possible expression of a more elaborated visual representation of body shape for females than for males could be via the canonical + deviation scheme proposed here.

Recent work exploring the infants' understanding of the body's appearance is beginning to specify the cues to which infants are (in)sensitive, providing a foundation to test for asymmetric body representations developmentally (Bhatt et al., 2016; Slaughter et al., 2004; Zieber et al., 2014). In support of a developmental asymmetry in encoding human body sex, a recent eye tracking study (Alexander et al., 2016) reported that male and female infants aged between 3 and 18 months looked significantly longer at computer generated female figures. In this study, however, the authors did not directly test whether the sex of the infant's caregiver explained this preference leaving the expertise account still untested in body perception. Conversely, while we have substantial evidence on the adult's encoding of sex from facial shape and texture (Bruce et al., 1993; Hill et al., 1995; Kramer et al., 2017; Le Gal and Bruce, 2002), it is unknown whether a similar asymmetry describes those representations. Finally, previous evidence for "supramodal" encoding of sex, provided by cross-adaptation studies (e.g. adaptation from face to body (Palumbo et al., 2015); or from body to face (Ghuman et al., 2010) suggests that mental representations of others' sex are multilevel, encompassing both cue-specific and cue-general descriptions of how the sexes differ. Further search studies can reveal at which of these levels asymmetric representations are found.

In sum, the present findings expose the structure of mental representations that relate sex and body shape. In doing so, they improve our understanding of how observers efficiently use body cues to make inferences about others. More broadly, a

better understanding of body perception will support direct comparisons to models of how we perceive faces (Duchaine & Yovel, 2015) and voices (Latinus & Belin, 2011), which in turn will facilitate efforts to develop a framework for person perception and categorisation in general (Freeman & Ambady, 2011; Yovel & Belin, 2013).

CHAPTER 3

Perceiving emotion and sex from the body: evidence from the Garner task for independent processes

This chapter investigates the relation between socially relevant information conveyed by the human body. Specifically, in this set of studies, I implement social vision approaches by exploring whether an interaction between body sex and body emotional expression information is present on a perceptual level. To answer whether our attention system independently or integrally processes these two highly salient social cues, we use the Garner selective attention task over three experiments.

This chapter is published as it is from:

Gandolfo, M., & Downing, P. E. (2019). Perceiving emotion and sex from the body: evidence from the Garner task for independent processes. *Cognition and Emotion*, 1-11.

Acknowledgements:

We thank Simona Skripkauskaitė and Julia Landsiedel for helpful comments on the manuscript.

3.1 Abstract

The appearance of the body signals socially relevant states and traits, but how these cues are perceived is not well understood. Here we examined judgments of emotion and sex from the body's appearance. Understanding how we extract these cues is important because they are both salient and socially relevant. Participants viewed body images and either reported the emotion expressed by each body while ignoring its sex, or else reported the sex while ignoring its emotion. Following Garner's logic (1974), two types of blocks were compared. In control blocks, the task-irrelevant dimension was fixed (e.g. all male in an emotion judgment task), whereas in orthogonal blocks it varied orthogonally to the task-relevant dimension (e.g. male-female). Where two dimensions draw on shared processes, interference results in relatively slower responses during orthogonal blocks. In contrast, a finding of no Garner interference –efficient selection of the task-relevant dimension– is taken to reflect independent processes. Bayesian analyses revealed evidence of no Garner interference between sex and emotion judgments, showing that extraction of these distinct signals from the body's appearance proceeds along largely parallel processing streams. These findings are informative about the mental architecture behind our perception of socially relevant characteristics of other people.

3.2 Introduction

From a mere glimpse, we make inferences about other people based on their visual appearance, and these inferences shape our social behaviour (Adams et al., 2011). To date, the major focus in the field of “social vision” has been on face perception, leading to a generally accepted model of face processing (Bruce and Young, 1986; Haxby et al., 2000; Duchaine and Yovel, 2015). There is also strong evidence that the visual appearance of the rest of the body provides a rich source of socially relevant cues (Aviezer et al., 2012; de Gelder, 2009; Johnson et al., 2012; Knoblich et al., 2006; Lawson et al., 2009; Rosenthal et al., 1979; Rice et al., 2013; Sell et al., 2009). Yet we remain some distance from establishing a “standard” model of body perception. Analogies from faces to bodies are likely to be imperfect given the significant perceptual differences between these two classes. However, the diversity of experimental approaches that has been used to carve up the processes of face perception can be extended to learn more about how we see bodies. In that spirit, here we used a selective-attention task to understand how representations of bodies are organised to support inferences about the sex and the emotional states of others.

There are strong a priori reasons to expect that body cues to sex should be highly salient for observers, given the clear biological and social relevance of this property. For example, adult observers are finely attuned to the shape differences that distinguish adult male from female body forms (Gaetano et al., 2014; Gandolfo and Downing, in review; Johnson and Tassinari, 2005) suggesting that this is a key dimension of the mental “spaces” that describe body shape (Palumbo et al., 2013). Further, discrimination of sex from static body images begins to develop in infants as early as 5 months of age (Hock et al., 2015) with continued refinement during childhood (Johnson et al., 2010). It is likewise understood that body cues to emotion are salient, given the high value of these signals to an observer (de Gelder, 2016). There is evidence for efficient extraction of emotional content from body postures (Martinez et al., 2016; Meeren et al., 2005) with a strong impact on spatial orienting as measured, for example, with saccade latencies (Bannerman et al., 2009) or by the influence of emotional body states on spatial neglect (Tamietto et al., 2007). Further, sensitivity to body shape and motion cues about emotions develops within the first few months of life (Bhatt et al., 2016). These kinds of evidence suggest that the need to perceive the sex and emotions of other people strongly influences the way bodies are mentally encoded.

Additionally, there are good reasons to expect that the processes behind perceiving sex and emotion may be intertwined. For one, at a perceptual level, there is evidence that the visual cues to sex and to emotion of faces are at least in part shared with each other (Becker et al., 2007; Bestelmeyer et al., 2010; Harris and Ciaramitaro, 2016; Hess et al., 2009) leading to behavioural interactions. For example, Bestelmeyer et al. (2010) found evidence for sex-contingent aftereffects following adaptation to facial emotion expressions of anger and fear. Furthermore, current dynamical views of social cognition and perception emphasize continuous interaction between bottom-up visual cues from social stimuli such as faces and bodies, and top-down social-cognitive processes related to categorical, associative, and contextual influences (Freeman and Ambady, 2011; Freeman and Johnson, 2016). In the specific case of sex and emotion, for example, strong cultural associations link femininity with a more intense experience and expression of emotion, and link masculinity with anger and pride (Becker et al., 2007; Kite et al., 2008; Plant et al., 2000). In line with an interactive view, then, these associations may in turn shape and influence “lower level” perceptual processes. Such interactions play out in evidence that observers’ evaluation of emotional expressions can be coloured by their understanding of the sex of the model (Condry and Condry, 1976; Haugh et al., 1980), and conversely, that emotional facial expressions bias the evaluation of sex (Bayet et al., 2015). More recently, a similar interaction has been observed for body perception, in that emotional expression judgments are biased by the sex of the model, often in a stereotypical direction (e.g. female = positive emotion; Bijlstra et al., 2018).

Taken together, the preceding considerations motivated the present study, which examined whether visual representations of sex and emotion from the body are independent to each other – proceeding along different processing “channels” – or rather instead are functionally integrated. Specifically, we adapted the Garner selective attention task, which has provided an important tool for establishing the independence or otherwise of perceptual processes (Algom and Fitousi, 2016; Garner, 1974). In such tasks, participants are required to classify exemplars of a stimulus along a particular dimension (e.g. shape) while a second, task-irrelevant dimension of the stimulus (e.g. colour) is either held constant or is varied. Performance (in terms of response times) is typically compared in two key conditions. In a “control” condition, the irrelevant dimension is kept stable (e.g. squares and rectangles all shown in red). In an

“orthogonal” condition, the irrelevant dimension varies orthogonally across trials (e.g. squares and rectangles in either red or blue). A relative increase in response times in the orthogonal condition indicates that participants cannot fully filter out the irrelevant dimension (in this example, colour) while attending to the shape information. The logic in this case is that perception of the two dimensions is at least partly integral – that is, reliant on shared perceptual processes. Conversely, if no impairment in the orthogonal condition is found, the two dimensions of the stimulus are considered to be processed independently.

The Garner logic has been applied to test perception of properties carried by the appearance of the face such as sex, identity and expression (Amishav and Kimchi, 2010; Ganel and Goshen-Gottstein 2002; 2004; Schweinberger and Soukoup, 1998; Schweinberger et al., 1999). For example, Ganel and Goshen-Gottstein (2002) found evidence for integral processing of face identity and sex, in support of a shared- rather than independent-route hypothesis. Such findings demonstrate the utility of selective attention tasks to test the processing architecture of complex, meaningful stimuli in addition to the more elemental perceptual dimensions tested in earlier work.

More recently, Johnstone and Downing (2017; see also Reed et al., 2018) reported the first application of the Garner approach to body perception. They investigated two socially-relevant properties that are conveyed by body shape: sex and weight. They found that irrelevant variation of sex interfered with weight judgments, but irrelevant variation of weight did not interfere with sex judgments. One interpretation of this asymmetrical pattern of interference is that it reflects a parallel contingent system (cf. Turvey, 1973). That is, parallel processes analyse features of body shape related to sex and to weight, and initial results of sex categorisation impact the processing of weight (but not vice versa). This architecture may relate to the distinct impacts of weight gain/loss on body shape for males and for females. These findings illustrate the potential of the Garner approach for dissecting the visual perception of the body.

In the present study, we adopted Garner’s approach to examine the perception of sex and emotional states from others’ bodies. The emotional expressions we tested were drawn from some of the canonical emotion categories that have been extensively explored in studies of face and body emotional expression: anger, fearful, happy, sad. These were paired arbitrarily so that each emotion task would require a binary judgment

comparable to the sex task. Hence over three experiments, we asked participants to perform a sex judgment (male/female) and, in separate blocks, an emotion judgment (Experiment 1: angry/fearful; Experiment 2: angry/happy; Experiment 3: happy/sad). The judgments were performed on images of real human bodies, which, crucially, were the same in both tasks. In line with the Garner logic, participants performed the tasks in Control blocks (in which the irrelevant dimension was held constant) and in Orthogonal blocks (in which the irrelevant dimension was varied orthogonally). If accurate response times are found to be slower in the Orthogonal condition, this would provide evidence for shared representations underlying the extraction of sex and emotion from the appearance of the body. In contrast, if performance is unhindered by variation in the irrelevant dimensions, this indicates that these social cues are perceptually extracted by independent processes.

3.3 Methods

3.3.1 Participants

Participants in each experiment were 32 students at Bangor University. Sample size was informally determined *a priori* with reference to similar previous studies (e.g. Johnstone and Downing, 2017). No participant took part in more than one of the experiments. They took part in return for course credit in a research methods module, and provided informed consent to participate. Gender ratios and age ranges are reported below with the results of each experiment. The experimental procedures were approved by the Research Ethics Committee of Bangor University's School of Psychology.

3.3.2 Stimuli

Digital photographs of 8 different actors (4 males, 4 females) performing 4 different emotional expressions (anger, fear, happy, sad) were captured against a neutral background. Each actor wore a white t-shirt and blue jeans. Actors were instructed to express the emotions using their whole body and to avoid raising their hands and arms above the head or in front of the face. They were not discouraged from expressing the emotion with the face, although they were aware that the face would be obscured in the resulting image. Images were re-scaled from their original raw size to

be 450 x 450 pixels (approximately 13 x 13 cm onscreen) and converted to grayscale. In each image, the face and head were covered with a mosaic mask using Photoshop (Adobe Inc.). Feet were cropped from the images given that (unlike clothing) shoe type and colour were not held constant across models. A total of 148 images were collected in the initial stimulus-creation phase. Using Google Forms (Google Inc.), we then collected online ratings data from 25 participants regarding these images. These participants were asked to categorise which emotion was expressed in each image, by choosing between “Angry”, “Happy”, “Sad”, “Fear” or “Other”. They were also asked for each image how intense and how realistic the emotion expression was, on a Likert scale from 1 to 5. For each emotion and each actor, we selected the most accurately identified images, with the result that all selected images were classified with above 90% mean accuracy. We then selected, for each emotion x sex x actor combination, the image with the best combined rating for realism and intensity. The final set of stimuli comprised 8 pictures for each emotion. Sample images are shown in Figure 3.1. The full set is shown in Supplementary Figure 1.

3.3.3 Design and Procedure

Two speeded binary classification tasks were performed by each participant. In each experiment, 16 participants first performed an emotion judgement task (for example, Anger vs Fear), and 16 performed a sex judgement task first. Task order was assigned on the basis of registration for the study. To counteract carryover effects, for each task, the Control and Orthogonal blocks were presented in a counterbalanced order across participants. Two versions of each Control condition were tested, to include trials for each level of the irrelevant dimension. For example, in the Angry vs Fear task, in one Control block the images were all of females, and in the other all males. These blocks were presented in counterbalanced order across participants. Continuing the example, in the Orthogonal block, the images would be a mixture of males and females. Each Control block comprised 64 trials, and the Orthogonal block comprised 128 trials. Block structure was not made explicit to the participants, in order to avoid drawing their attention to changes in the irrelevant stimulus dimensions. Participants had a break between the two classification tasks.

The experiment was administered using the Psychtoolbox (Brainard, 1997) package running in Matlab (MATLAB Release 2012b, The MathWorks, Inc.) on an Apple iMac computer. Viewing distance was approximately 60 cm from the screen but was not fixed. In each trial, following a 300 ms fixation dot, a single stimulus was presented at the centre of the screen, where it remained until the participant responded or until a maximum duration of 1.5 sec. Participants' responses were recorded with the "f" and "j" keyboard keys to report either the sex or the emotion of the body. They were reminded of the response mapping with the corresponding category names printed at the bottom of the screen. Participants were instructed to respond as quickly and as accurately as possible.

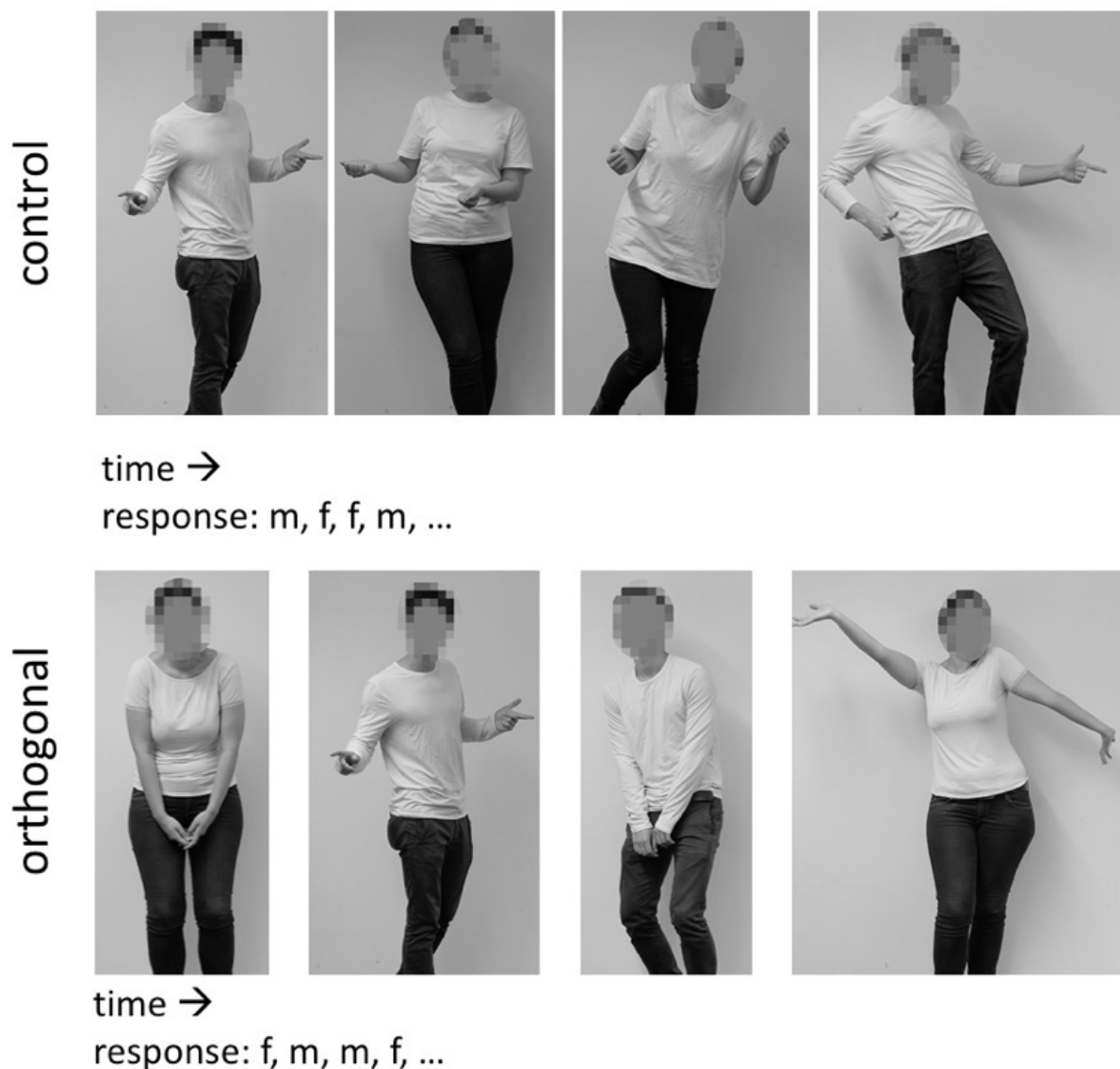


Figure 3.1. Schematic illustration of the Garner task as applied in the present study. In a Control block, participants make one binary judgment (here, on sex) on bodies presented individually, whilst the other irrelevant dimension is held constant (here, happy emotional body expression).

In an Orthogonal block, the irrelevant dimension varies (here, between happy and sad emotional body expressions). If response times are slower in the Orthogonal blocks, this shows that filtering of the irrelevant dimension has (at least partially) failed, which would indicate shared processing routes for the dimensions of sex and emotion. In contrast, where performance in the two blocks is equivalent, this indicates successful filtering, enabled by parallel underpinning processes.

3.3.4 Analyses

Two separate ANOVAs were conducted on the response time and accuracy data, respectively, from each experiment: these took the form of 2 (Task: Sex or Emotion) X 2 (Block: Control or Orthogonal) repeated measures factorials. For the response time analysis, only data from accurate trials were included. Data from participants whose overall accuracy was < 2.5 SD below the group mean, or whose mean response times were >2.5 SD above or below the group mean, were removed and new participants were tested to bring the N up to 32.

Following the logic of the Garner selective attention task, the absence of interference (no difference in response times between Orthogonal and Control blocks) can be considered as positive evidence for independent processing of the two dimensions in question. Accordingly, here we complemented the null hypothesis significance testing approach with a Bayesian inference approach in order to assess to what extent our findings evidence the independence of processing sex and emotion from the body. To do this, we computed Garner interference indexes for each participant (mean accurate response time for Orthogonal – mean accurate response time for Control condition) and performed one sample t-tests against 0 (i.e. absence of Garner interference). Bayes Factors (BF) were computed over these values using Jasp (Jasp software; <https://jasp-stats.org/>) (Love et al., 2015). We considered findings of BF_{10} lower than $1/3$ (Jeffrys, 1961; Lee and Wagenmakers, 2014) as moderate or better evidence for independent processing of the two dimensions. Results for all three experiments are shown in Figure 3.2 and provided numerically in Table 3.1.

3.4 Results

3.4.1 Experiment 1: Angry and Fearful

Participants in Experiment 1 consisted of 27 females and 5 males, with a mean age of 19 ± 2.2 years. Data from 1 participant were excluded from the analysis due to poor performance and were replaced.

The ANOVA on accuracy did not reveal any significant effects (all p s > 0.11). The ANOVA on response times showed a significant main effect of Task, $F(1, 31) = 23.41$, $p < 0.001$, $\eta^2 = 0.43$. Discrimination of sex was faster ($M = 621$ ms, $SE = 14.1$) than of emotional postures ($M = 669$ ms, $SE = 14.5$). No other effect reached significance (all $p > 0.55$).

A Bayesian one sample t-test on Garner interference indexes showed moderate evidence for the absence of Garner interference both in the sex task ($BF_{10} = 0.21$) and in the emotion task ($BF_{10} = 0.21$) in the sense that the null hypothesis is at least three times more likely than the alternative.

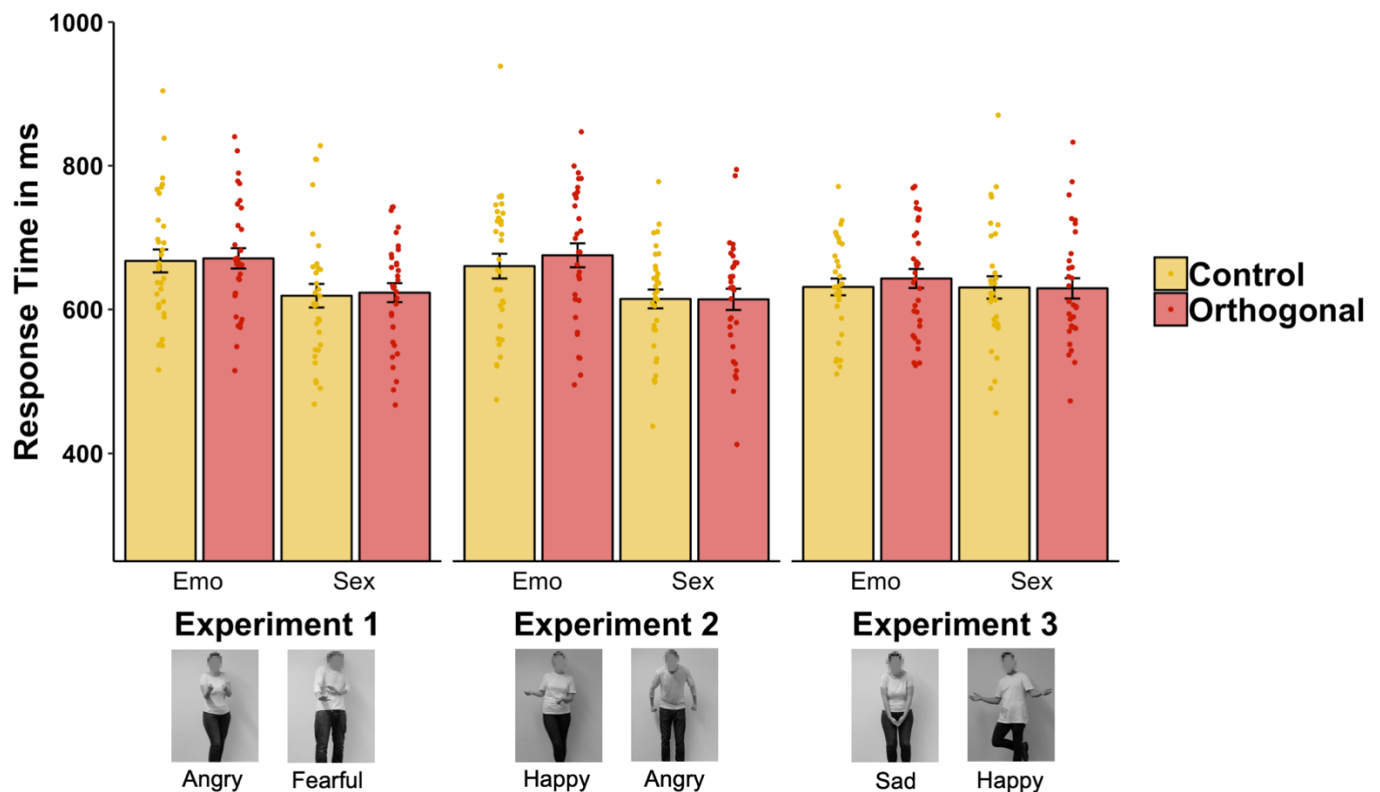


Figure 3.2. Accurate response times in each experiment and condition. Bars represent the mean across participants; error bars reflect SE of the mean including both within- and between-

participants variance; each point reflects one participant's data. Example stimuli from each emotion category are shown at bottom. All experimental stimuli are provided in Supplementary Figure S3.1.

		RTs in ms				Accuracy			
		EMO- Control	EMO- Orthogonal	SEX- Control	SEX- Orthogonal	EMO- Control	EMO- Orthogonal	SEX- Control	SEX- Orthogonal
Exp. 1	<i>M</i>	667.57	671.13	619.07	623.33	0.93	0.94	0.91	0.92
	<i>SEM</i>	15.93	14.13	16.38	13.27	0.008	0.006	0.009	0.009
Exp. 2	<i>M</i>	660.37	675.38	614.64	614.07	0.87	0.87	0.93	0.92
	<i>SEM</i>	17.24	16.71	13.09	14.78	0.01	0.01	0.01	0.01
Exp. 3	<i>M</i>	631.35	643.09	630.6	629.35	0.93	0.93	0.91	0.92
	<i>SEM</i>	11.69	13.3	15.59	14.15	0.009	0.008	0.01	0.01

Table 3.1. Mean response times (with SEM) and accuracy, separately for each condition and for each experiment.

3.4.2 Experiment 2: Happy and Angry

Participants in Experiment 2 consisted of 26 females and 6 males, with a mean age of 19 ± 1.1 years. Data from 3 participants were excluded from the analysis due to poor performance and were replaced.

The ANOVA on accuracy showed a main effect of Task, $F(1, 31) = 9.92$, $p < 0.001$, $\eta^2 = 0.46$. Participants were more accurate in classifying sex ($M = 0.93$, $SE = 0.01$) than emotion ($M = 0.87$, $SE = 0.009$). No other effect reached significance (all $p > 0.66$).

The ANOVA on response times showed a main effect of Task, $F(1, 31) = 32.27$, $p < 0.001$, $\eta^2 = 0.51$. Participants were faster in classifying sex ($M = 609$ ms, $SE = 12.1$) than emotion ($M = 668$ ms, $SE = 16.4$). No other effect reached significance (all $ps > 0.17$).

A Bayesian one sample t-test on Garner interference indexes showed moderate evidence in the sex task ($BF_{10} = 0.19$) and anecdotal evidence in the emotion task ($BF_{10} = 0.63$) for the absence of Garner interference.

3.4.3 Experiment 3: Happy and Sad

Participants in Experiment 3 consisted of 28 females and 4 males, with a mean age of 19 ± 1 years. Data from 2 participants were excluded from the analysis due to poor performance and were replaced.

The ANOVA on accuracy did not reveal any significant effects (all $p > 0.15$). The ANOVA on response times likewise did not reveal any significant effects (all $p > 0.38$).

A Bayesian one sample t-test on Garner Interference indexes showed moderate evidence for the absence of Garner interference both in the sex task ($BF_{10} = 0.19$) and in the emotion task ($BF_{10} = 0.33$) in the sense that the null hypothesis is at least three times more likely than the alternative.

3.5 General Discussion

Over three experiments, participants were able to successfully filter irrelevant variation in sex when judging bodily emotion, and *vice versa*. Broadly, this finding held across three different pairings of four different emotion categories. We take this as evidence for independent processes that support perceiving sex and emotion from body shape and posture.

Sex is generally considered to be one of the primary dimensions that perceivers establish upon encountering another individual (Stangor et al., 1992). Yet in the present study, participants were able to ignore this dimension successfully, with no detectable cost to performance on another body-related task. Some trivial accounts of this finding can be excluded. While the body images were constructed so that extraneous cues (face, clothing type) were obscured or controlled, this did not render the sex judgment task particularly difficult, as evidenced by the response time and accuracy data. This suggests that the absence of Garner interference from sex on the emotion task is not due to variation between sexes being minimized. Further, as the images were presented individually in the absence of other visual distractors, there was little perceptual load that might have interfered with extracting sex-related cues (Lavie, 1995).

The logic of interpreting interference effects depends on the relative baseline difficulty of the two discriminations being approximately equivalent. In the case that one task is much more difficult than the other, it can be trivial that the more difficult dimension does not interfere with the easier one. In the present study, overall task difficulty was not always perfectly matched. In Experiment 1, participants' discrimination of sex was faster (but no more accurate) than for emotion; in Experiment 2, the sex task was faster and more accurate. Note, however, that an easier sex task would predict that this dimension is more readily encoded than emotion, and hence if anything more likely to create interference in the Orthogonal condition. Yet no such interference was observed in those experiments, nor in Experiment 3 in which task difficulty was matched by both accuracy and response time measures.

As for sex, previous evidence has been mustered to argue that the body is a key signaller of emotions (Aviezer et al., 2012; de Gelder, 2016), and that such signals exert a powerful force on early perceptual processes (e.g. Tamietto et al., 2007; 2009). Yet in the present study, irrelevant variation in bodily emotion did not interfere with making sex

judgements. One possibility is that the stimuli did not strongly convey the emotions that they were designed to convey, and/or did not do so realistically. Our stimulus design procedure and our results provide assurance against the first concern. Performance on the emotion tasks was fast and accurate, showing that participants were readily able to make these discriminations. Further, the images were selected based on independent observers' ratings of the strength and clarity of the emotional expression they conveyed. Whether these instructed poses are *realistic* expressions of emotion – whether in daily life people spontaneously take these postures when happy, sad, etc. – is a separate concern, one that applies to all of the previous work on emotional body expression that uses posed expressions as stimuli. We do note that the stimuli used here were static images rather than dynamic displays, and it is possible that the salience or psychological validity of the emotions would be greater in movies (Peelen et al., 2007; Pichon et al., 2008).

As we used real human actors to generate our stimuli (as opposed to rendered images), we cannot exclude the possibility that male and female models differed systematically in the specifics of how they expressed a given emotion through posture. In other words, it is possible that the two dimensions tested here were in some sense entangled *in the stimulus*. However, such a confounding factor, if present and detectable by participants, would have been expected to produce Garner interference, which we did not observe.

Having excluded more trivial accounts of the present findings, it remains the case that arriving at a conclusion of independent processes depends on interpreting a null effect (that is, no difference between Control and Orthogonal blocks). This is a common feature of any application of the Garner paradigm. To this end, we applied Bayesian analyses in order to arrive at positive statements about the likelihood, given the data, of no Garner interference being present. Taken together, the present evidence shows that if Garner interference between sex and emotion in bodies exists, the true effect size is likely to be small.

How do the present findings compare with evidence on face perception? Previous studies have compared sex and emotion in faces with the Garner task, with varying results. Le Gal and Bruce (2002) found that angry / surprised facial expressions and sex were independently processed (i.e. no Garner interference) when the discriminability of the sex and emotion judgements were matched. In contrast, Atkinson et al. (2005) found an asymmetric interference pattern: the irrelevant variation of sex

interfered with happiness / fear expression judgments, but irrelevant variation of emotion did not interfere with sex judgements. They interpreted their finding by proposing that invariant dimensions of a person (sex) are more useful referents for computing information about variant aspects (emotional expression) than *vice versa*. More recently, Becker (2017) found asymmetric interference between emotion (anger, happiness) and sex, that further depended on individual differences in attentional control capacities. In future work, it would be useful to directly compare selective attention for faces and bodies in the same participants with matched designs and stimuli (including emotion categories), to better establish the commonalities and distinctions across these domains. Interestingly, the Garner approach has also provided evidence for the impact of mental health on the ability to attend or ignore emotional content. Gilboa-Schechtman et al. (2004) found that participants with depression were less able than controls to ignore emotional face content, relative to sex, in a Garner task.

How does the evidence for independent analysis of sex and emotion from the body relate to our understanding of the relevant neural systems? The last few decades have seen much new evidence on the location and properties of regions across the brain that are engaged in some way by the appearance of bodies and their movements (Brandman and Yovel, 2014; de Gelder 2006; Downing and Peelen, 2011, 2016; Minnebusch and Daum, 2009; Peelen and Downing, 2007; Schwoebel and Coslett, 2005). Whilst there has been significant interest in how these regions encode emotional expressions of the body (de Gelder, 2009; Peelen et al., 2007), little direct evidence exists on how they are involved in distinguishing the sexes. Nonetheless we can speculate about three possible scenarios linking brain and behaviour.

Two scenarios focus on cortical pathways. One view draws a distinction between the focal and selective brain representations of the body found in the ventral and dorsal temporal lobe regions. Ventrally, the extrastriate body area (EBA: Downing et al., 2001) and fusiform body area (FBA: Schwarzlose et al., 2005; Peelen and Downing, 2005) are engaged by static aspects of the body (e.g. shape), while dorsal regions (e.g. posterior superior temporal sulcus, pSTS) are engaged by dynamic displays of movements and meaningful actions (Downing et al., 2006; Giese and Poggio, 2003; Urgesi et al., 2007). Whilst the stimuli tested here were static, the emotional expressions could invoke a dynamic representation (e.g. by association with known movement patterns). In this case, we might relate the two independent processes identified here to engagement of the ventral (sex) and dorsal (emotion) pathways.

A second cortical view focuses within the ventral body-selective regions. Downing and Peelen (2011) surveyed neuroimaging findings to conclude that EBA and FBA jointly encode the shape and the posture of viewed bodies. Part of this conclusion was supported by evidence for encoding of emotion from body postures in the activity of EBA and FBA (Peelen et al., 2007; Atkinson et al., 2012; but see van de Riet et al., 2009). In this light, the present findings could reflect parallel processes that are spatially co-localised within these ventral regions, that represent our implicit knowledge of both the shape “space” that characterises different sexes, and also the postural representations that relate to different emotional states.

In contrast to the above perspectives, a third possible account (cf. de Gelder, 2006) emphasises the role of subcortical regions such as the amygdala in a rapid, automatic evaluation of bodily emotions, in particular those that might naturally evoke an overt response on the part of the observer (such as a defensive response elicited by fear). On this view, a subcortical process may work in tandem with cortical body-selective regions where a richer visual analysis of bodily expression is required, such as where finer verbalizable judgments about the emotions being expressed are needed. If this hypothesised pathway uniquely analyses emotion and does not extend to include processing of sex from body shape, then it offers another means by which these two body dimensions may be analysed independently. Adapting the Garner logic developed here to a functional neuroimaging design may help distinguish amongst these alternative accounts.

Finally, one limitation of the present study is that our sample was biased strongly in favour of female over male participants, precluding a systematic examination of the effect of participant gender on attention to sex and emotion. While there is not evidence to suggest an impact of observer gender on the independence of these two dimensions *per se*, there is active research on the influence of gender on the perception of socially relevant cues from faces and from bodies alike, with both positive and negative findings (e.g. Grimshaw et al., 2004; Kret and de Gelder, 2012; McBain et al., 2009; Montagne et al., 2005). Accordingly, future examinations of social cue processing from bodies should balance, and test for interactions with, participant gender.

3.6 Concluding remarks

While the discipline of social vision remains largely focused on face perception, there is increasing awareness of the need to better understand body perception. More

Towards a model of human body perception

broadly, a fuller understanding of person perception (from faces, bodies, and voices alike) contributes to the effort to draw together insights from perception with those from social psychology for a multi-level understanding of how we understand and interact with the social world (Freeman and Ambady, 2011).

CHAPTER 4

Causal evidence for Expression of Perceptual Expectations in Category-selective Extrastriate Regions

In this chapter, using fMRI-guided TMS, I investigate whether content-specific expectations towards a visual target, in the form of a verbal cue, are directly related to pre-activation of relevant neural populations involved in visually encoding that target. Importantly, here, expectations formed to an incoming human body picture are related to socially relevant information linked to its shape (i.e. sex).

This chapter is published as it is from:

Gandolfo, M., & Downing, P. E. (2019). Causal evidence for expression of perceptual expectations in category-selective extrastriate regions. *Current Biology*, 29(15), 2496-2500.

Author Contributions:

M.G. and P.E.D. conceived and designed the experiments. Data collection was performed by M.G. M.G. and P.E.D. analyzed the data. The manuscript was written by P.E.D and M.G.

Acknowledgements:

We thank M. Peelen, C. Urgesi, K. Valyear, and the Bangor Imaging Group for valuable feedback and advice. L. Jacobs assisted with task development and data collection, and K. Darda with data collection. M.G. was funded by a PhD student grant from Bangor University.

4.1 Summary

Expectations about a visual event shape the way it is perceived (Bruner and Postman, 1949; de Lange et al., 2018; Lupyan and Clark, 2015; Summerfield and Egner, 2009). For example, expectations induced by valid cues signalling aspects of a visual target can improve judgments about that target, relative to invalid cues (Battistoni et al., 2017; Carrasco, 2011). Such expectation effects are thought to arise via pre-activation of a template in neural populations that represent the target (Simanova et al., 2016; Summerfield and de Lange, 2014) in early sensory areas (Kok et al., 2014) or in higher-level regions. For example, category cues (“face” or “house”) modulate pre-target functional MRI (fMRI) activity in associated category-selective brain regions (Esterman and Yantis, 2010; Puri et al., 2009). Further, a relationship is sometimes found between the strength of template activity, and success in perceptual tasks on the target (Kok et al., 2017; Peelen and Kastner, 2011; Stokes et al., 2009). However, causal evidence linking pre-target activity with expectation effects is lacking. Here we provide such evidence, using fMRI-guided online transcranial magnetic stimulation (TMS). In two experiments, human volunteers made binary judgments about images of either a body or a scene. Before each target image, a verbal cue validly or invalidly indicated a property of the image, thus creating perceptual expectations about it. To disrupt these expectations, we stimulated category selective visual brain regions (extrastriate body area, EBA; occipital place area, OPA) during the presentation of the cue. Stimulation ended before the target images appeared. We found a double dissociation: TMS to EBA during the cue period removed validity effects only in the body task, while stimulating OPA removed validity effects only in the scene task. Perceptual expectations are expressed by the selective activation of relevant populations within brain regions that encode the target.

Keywords: perceptual expectations; pre-stimulus brain activity; category-selective brain regions; transcranial magnetic stimulation; extrastriate body area; occipital place area

4.2 Results and discussion

We designed two visual tasks that demonstrate the effects of verbal cues on the efficiency of perceptual judgments, and that are aligned to suitable cortical targets for brain stimulation (Figure 4.1A). In a body perception task, valid verbal cues about the sex of a target body image (“m” or “f”) improved the efficiency (mean RT / p(correct)) of judgments about the weight of the depicted person (heavy vs slim), relative to invalid cues, $t(24)=2.43$, $p=0.02$, $d=0.49$ (cf. (Johnstone and Downing, 2017)). In a scene perception task, valid verbal cues about the semantic category of a scene (“kitchen” or “garden”) improved the efficiency of judgments about the target image’s orientation (upright vs inverted; cf. (Lupyan and Thompson-Schill, 2012), relative to invalid cues, $t(24)=2.81$, $p=0.02$, $d=0.56$.

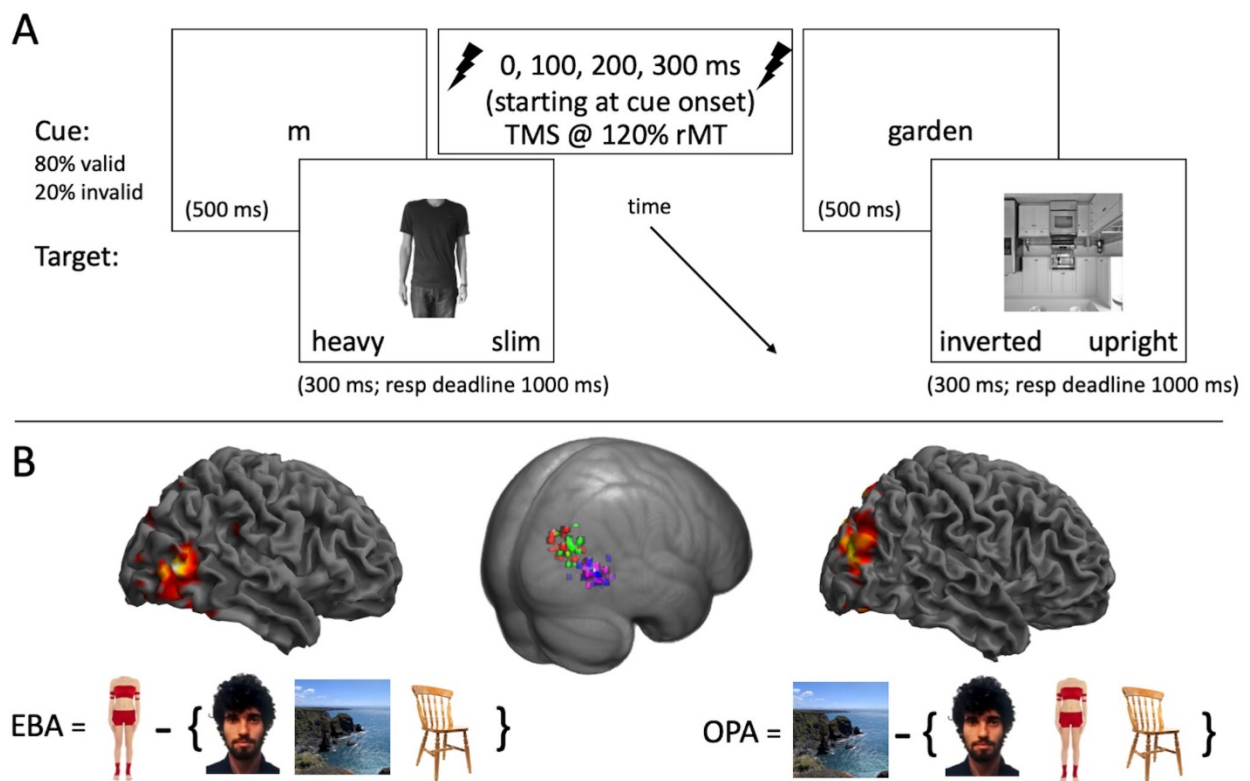


Figure 4.1. Schematic illustration of task timeline and targeted brain regions. (A) Timeline of the body task (left) and the scene task (right). In each case, a written cue predicted, with 80% validity, a property of the target image that next appeared. Participants made a binary weight judgment on each body image (heavy vs slim) or a binary orientation judgment on each scene image (inverted vs upright). Display images not to scale. (B) fMRI-guided transcranial magnetic stimulation (TMS) was used to interrupt activity in body (left) and scene (right) selective occipitotemporal brain regions during the processing of the cues, and before the onset of the target images. Activation maps on gray-matter surfaces show representative localisation of extrastriate body area (EBA; left) and occipital place area (OPA; right). Point clouds on brain

surface (centre) show, in MNI space, targeted peak locations for each participant x task combination (see also Table S1). Red: OPA, scene task; green: OPA, body task; blue: EBA, scene task; pink: EBA, body task.

We then used online transcranial magnetic stimulation (TMS) in two experiments testing these two tasks, to establish that cue-driven neural activity in category-selective occipitotemporal regions is causally necessary for the expression of these validity effects. Neuroimaging studies have identified focal regions of occipitotemporal cortex that are selectively involved in body and scene perception. The activity of these regions relates to online visual representation of their preferred categories – for example, encoding body shape and posture (Downing and Peelen, 2011; Zimmermann et al., 2016) in the extrastriate body area (EBA), and describing scene geometry (Dillon et al., 2018; Julian et al., 2016) in the occipital place area (OPA). Further, TMS studies have demonstrated a category-selective causal role for these regions in visual detection and discrimination tasks (Dilks et al., 2013; Downing and Peelen, 2016; Pitcher et al., 2009; Pitcher et al., 2012; Urgesi et al., 2004; van Koningsbruggen et al., 2013).

Using fMRI-guided TMS applied online during cue presentation (and ending before the target image appeared; Figure 4.1B; see also Table S1) we found that expectations in the body and scene tasks were selectively instantiated by activity in EBA and OPA respectively. In the first TMS study (N=21), a significant validity effect was found in the body perception task when TMS was applied to OPA ($t=2.14$, $p=0.045$, $d = 0.47$) but not to EBA ($t=-1.47$, $p=0.16$, $d = -0.32$; Site x Validity, $F(1,20)=5.3$, $p = 0.032$, $\eta^2=0.21$). In the second study (pre-registered; N=21), an effect of cue validity was found in the scene task when TMS was applied to EBA ($t=2.70$, $p=0.013$, $d = 0.59$) but not to OPA ($t=0.57$, $p=0.57$, $d = 0.12$; Site x Validity, $F(1,20)=6.1$, $p=0.023$, $\eta^2=0.23$). Direct comparison of the two studies shows that the influence of cues on efficiency was disrupted in a task- and region-specific fashion (interaction of stimulation Site x Task x Validity in a mixed-design ANOVA, $F(1,40) = 11.34$, $p = 0.00017$, $\eta^2=0.22$). Collapsing over the two studies shows that when TMS was applied to the task-relevant regions (EBA for bodies, OPA for scenes), cue validity effects were on average eliminated ($M = -1$ ms, $t=-0.15$, $p=0.87$, $d = -0.02$) while they remained significant when TMS was applied to the task-irrelevant regions (EBA for scenes, OPA for bodies) ($M = 26$ ms, $t=3.33$, $p=0.0018$, $d = 0.51$) (Figure 4.2; see also Figure S4.1

and Table S4.2). Pre-target activity in category-selective regions is causally necessary to express the perceptual expectations generated by verbal cues.

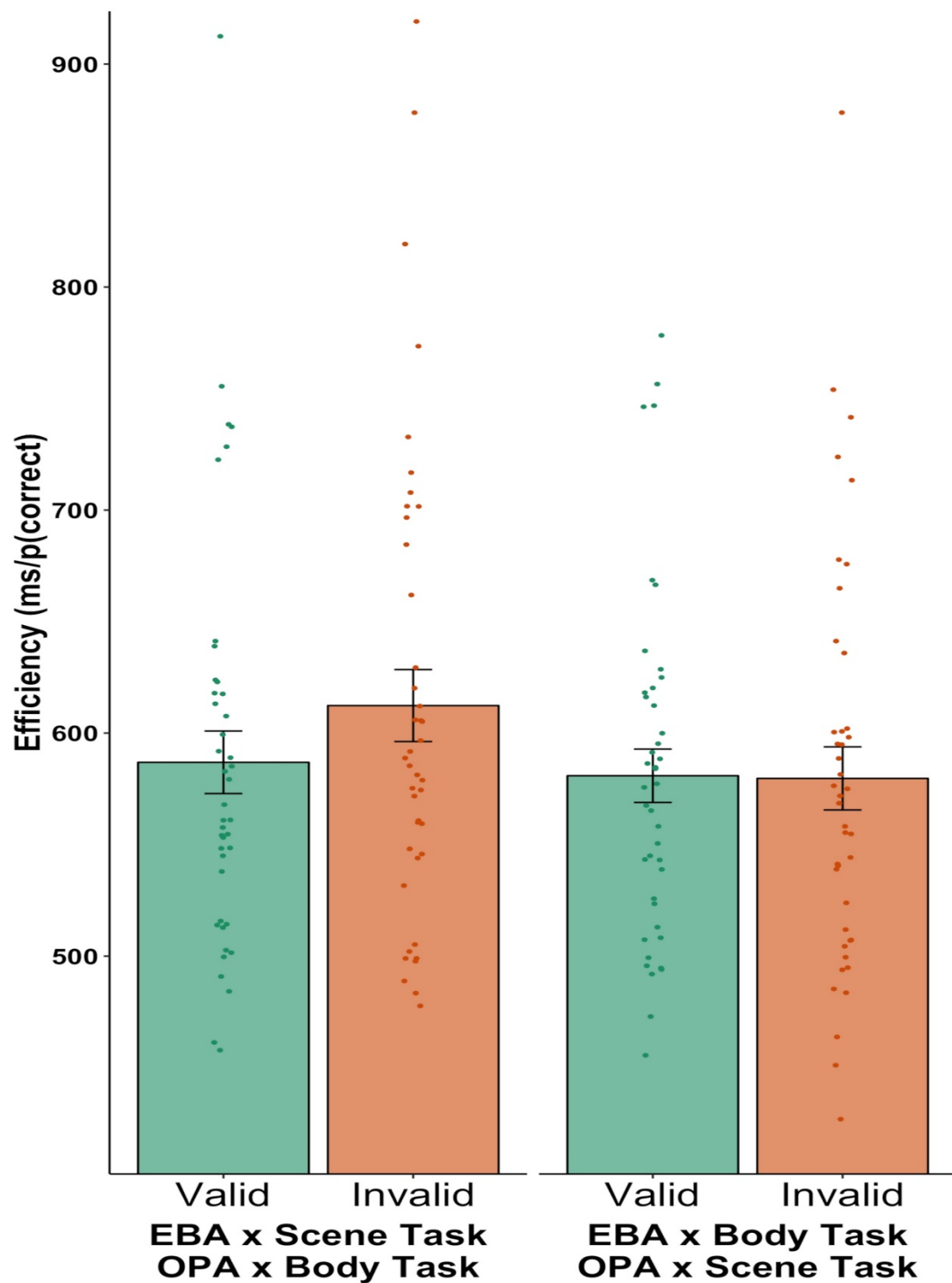


Figure 4.2. Impact of TMS over extrastriate category-selective regions on cue-driven stimulus expectations. Mean efficiency scores across participants (RT / p(correct)) are plotted in relation to cue validity, separately for conditions in which TMS was applied to the task-irrelevant (left)

and the task-relevant (right) brain regions. TMS during the cueing interval selectively eliminated effects of cue validity when applied to the task-relevant sites. Bars indicate mean values; error bars SE of the mean (including within- and between-participants variance); individual points reflect scores for each participant. See also Figure S4.1 and Table S4.2.

We speculate that the expectancy effects revealed (and disrupted) here relate to domain-specific aspects of the structure of body and scene encoding in EBA and OPA respectively. For example, sex reflects a core division within visual body representations, due to its relevance over evolutionary and lifetime scales (Hock et al., 2015; Johnson and Tassinari, 2005). In turn, the representation of each sex can be characterised by distinct mental “spaces” that capture the relationships between body shape and weight (Hill et al., 2016). On this view, expectations in the body task are reflected in the selection of subsets of the spaces describing likely body shapes of the cued sex. This hypothetical selection process can be construed as a form of internal attention (Chun et al., 2011) or as a form of neural sharpening, as described in previous studies of expectancy effects in vision and action (Kok et al., 2012; Yon et al., 2018).

Turning to scenes, images of different environments differ in their visual properties, in the kinds and distribution of objects present, and in the boundaries and distances implied (Greene, 2013; Malcolm et al., 2016). These considerations suggest two mechanisms by which expectancies (“kitchen” vs “garden”) could facilitate judgments of scene orientation. First, scene gist enhances localisation and identification of expected objects (Biederman et al., 1982; Davenport and Potter, 2004; Peelen and Kastner, 2014), and such objects may in turn support a scene orientation judgment. Second, different environments differ in openness and in the number and nature of their boundaries (Oliva and Torralba, 2001): while garden scenes tend to be open and contain fewer navigationally-relevant boundaries, indoor scenes such as kitchens are generally enclosed and more constrained. Such regularities may help to select the areas within a scene image that are diagnostic of its orientation. These proposals are consistent with evidence that OPA plays a role in encoding objects (Kamps et al., 2016; Troiani et al., 2014) and scene boundaries (Julian et al., 2016).

Our TMS findings are specific to the combination of stimulation site, task, and validity, ruling out several potential confounds. For example, these selective effects

cannot be explained by distracting effects of peripheral muscle stimulation, by a general alerting effect of the cues, or by disruption of general linguistic processes related to reading those cues. Further, our study is better controlled than those that compare stimulation over an active site to sham stimulation, or to the vertex (Meteyard and Holmes, 2018), because we stimulated two functionally comparable and adjacent (mean Euclidean distance between targeted peaks = 2.55 cm) sites. (Secondary planned ANCOVAs showed no evidence for a systematic inter-participant relationship between the distance between sites and the Stimulation Site x Validity interaction effects: body task: $p = 0.26$; scene task: $p = 0.28$).

Several lines of evidence suggest that the effects of stimulation were more likely related to processes triggered by the verbal cues than to spillover of TMS effects directly onto online visual perception of the targets. One line of evidence relates to the timing of relevant neural activity. Previous work showed that post-stimulus TMS over EBA is more effective than pre-stimulus TMS at interfering with performance on a person detection task (van Koningsbruggen et al., 2013). Further, the earliest category-selective effects of TMS over EBA on a visual discrimination task are found over a narrow temporal window around 100-110 ms after stimulus onset, roughly 300 ms after the final pulse in our protocol (Pitcher et al., 2012). While equivalent TMS data are not available for OPA, a recent magneto-encephalography study showed that a texture-independent representation of scene geometry likewise first emerges in this region at about 100 ms after stimulus onset (Henriksson et al., 2019). Moreover, setting aside timing considerations, if TMS were directly impacting stimulus-driven perceptual processes, then we would expect overall performance to be impaired (collapsing over valid and invalid conditions) when task-relevant regions were stimulated, compared to task-irrelevant regions. In fact, in each experiment the non-significant trend was in the opposite direction (main effect of Stimulation Site: body task, EBA: 584 ms, OPA: 592 ms, $p=0.24$; scene task, EBA: 607, OPA: 577 ms, $p=0.08$). Alongside the significant Site x Task x Validity interaction, these findings strongly suggest that the main impact of TMS in this study is on cue-related expectation processes rather than directly on image perception per se.

Other aspects of these tasks allow us to specify the expectation-related processes that they capture. First, in both tasks the cues were orthogonal to the task-

relevant dimensions: they predicted which of two possible types the target would reflect, but not which response would be required. As such, the cue-related neural activity in EBA and OPA must have been related to forming expectations about the target itself, rather than about the decision or response required. Second, because the cues were in a different format than the targets, the effect of validity cannot be attributed to visual similarity between the cue and the target, and must instead have been at a more abstract level. Third, the targets in these tasks were presented in isolation and well above threshold. As such, the observed cueing effects were not related to filtering out distractors, or to consolidating awareness of ambiguous or near-threshold stimuli (cf. (Panichello et al., 2012; Summerfield et al., 2006)). Finally, these findings are not attributable to state-dependent effects of TMS (Ambrus et al., 2019; Bergmann, 2018; Silvanto et al., 2008): owing to the design counterbalancing, regional brain states at the time of brain stimulation were balanced, on average, with regard to the main validity manipulation.

Pre-stimulus brain activity in occipitotemporal regions is critical for the expression of perceptual expectations about those regions' preferred stimuli. This finding does not rule out additional causal contributions from other, domain general mechanisms. For example, selective attention may partly mediate the effects of an expectation on perception of the target. In the body task, for example, a sex cue may direct attention towards regions of the body that reliably distinguish heavy and slim people of that sex. As such, we see in these tasks an interplay between expectations generated by cues, and selection processes that facilitate turning those expectations into behavioural benefits. More broadly, forming perceptual expectations must also rely on flexible mechanisms that can interpret cues and relate them dynamically to current task goals. While the present findings do not speak to the neural basis of such mechanisms, a proposed hub-and-spoke network for controlled semantic cognition (Ralph et al., 2017) appears to have the requisite components to link the verbal, visual, and semantic properties of people and places as tested here.

4.3 Methods

The procedures were approved by the Research Ethics Committee of Bangor University's School of Psychology. Participants were students at Bangor University and provided informed consent for their participation. They took part in return for course credit in a research methods module, or for a cash payment. No individual participant took part in more than one experiment.

4.3.1 Participants: Behaviour-only

Fifty-four participants took part in two experiments. Twenty-seven of these participated in the body perception task (4 males; mean age 20 ± 3) and twenty-seven in the scene perception task (5 males; mean age 21 ± 6). Two participants from each task were excluded because their mean accuracy or response times were 2.5 or more SDs above or below the group mean across conditions for that task. The final sample comprised 25 participants in each task.

4.3.2 Participants: TMS

Forty-seven participants took part in the TMS experiments. They were screened following the safety screening standard questionnaire for rTMS (Rossi et al., 2009; Rossi et al., 2011). None of the participants reported any history of neurological, psychiatric or other major medical disorders. Twenty-three of these participants performed the body task (12 males; mean age: 24 ± 3 years) and twenty-four performed the scene task (6 males; mean age: 22 ± 5 years). One participant from the body perception task and one participant from the scene task were excluded because accuracy was 2.5 or more SDs below the group mean across conditions for that experiment. Three more participants (1 from the body perception task and 2 from the scene task) were excluded due to experimenter error or motion/discomfort during the stimulation. The final sample comprised 21 participants in each task. The sample size of the scene task was pre-registered to match the final sample of the body perception task together with the other experimental procedures (link: <http://aspredicted.org/blind.php?x=xu95zn>).

4.3.3 Imaging

Each participant in the TMS experiments first completed two to four runs of a four conditions block-design functional localiser fMRI experiment in order to identify target sites for stimulation. The stimuli consisted of blocks of images of human bodies (without heads), unfamiliar faces, outdoor scenes, and chairs. Each condition was presented in four blocks of 18 sec in each run. These were interspersed with 5 fixation blocks of 16 sec duration, resulting in a total of 21 blocks per run. In each block, 24 images (selected randomly from a full set of 40) were presented, each for 300 ms followed by a 450 ms blank interval. During each block, an image was presented twice in a row two times. Participants were instructed to detect these repetitions and press a key (1-back task).

Imaging data were acquired using a 3T Philips MRI scanner with a 32-channel SENSE phased-array head coil. Functional data (T2* weighted, gradient echo sequence; echo time, 35ms; flip angle, 90°) were acquired with the following scanning parameters: repetition time 2 seconds; 35 off-axial slices; voxel dimensions 3x3 mm; 3mm slice thickness; SENSE factor 2, phase encoding direction anterior-posterior. A high-resolution anatomical scan was also acquired (T1 weighted, 175 sagittally oriented slices; 1mm isotropic voxels; repetition time, 8.4 ms; echo time, 3.8ms; flip angle, 8°).

Functional MRI data were preprocessed and analysed using SPM12 (Wellcome Department of Imaging Neuroscience, London, www.fil.ion.ucl.ac.uk/spm/software/spm12/). The functional images were realigned and spatially smoothed (6-mm FWHM Gaussian kernel). The resulting images were entered into a subject-specific general linear model with four conditions of interest corresponding to the four categories of visual stimuli. Estimates of the BOLD response in each voxel and category were derived by a general linear model including the boxcar functions of stimulation that were convolved with a standard hemodynamic response function. All analyses were performed in participant-native coordinates; for reporting purposes, target sites were converted to standard MNI space.

In each participant individually, we localised right hemisphere body and scene selective regions by contrasting the response to human bodies with that to the remaining three conditions and the response to scenes with that to the remaining three conditions respectively. Each TMS target site (right hemisphere extrastriate body area [EBA]; right hemisphere occipital place area [OPA]) was individually identified by

selecting the peak activation for that category in the relevant lateral occipito-temporal region based on previous findings (Julian et al., 2016; van Koningsbruggen et al., 2013). The mean peak MNI coordinate (X, Y, Z, with SEs) was 48 (0.65), -71 (0.98), 2 (0.72) for right EBA and 34 (0.80), -79 (0.70), 20 (0.93) for right OPA (see also Table S4.1).

4.3.4 TMS stimulation

A Magstim Rapid2 (Magstim; Whitland, UK) with a 70mm figure-eight coil was used for the TMS. Stimulation intensity was set at 120% of the resting motor threshold, defined as the minimal intensity of left motor cortex stimulation required to elicit a reliable MEP of at least 50 μ V in the right hand's first dorsal interosseous muscle (Rossini et al., 2015). Online TMS was delivered at 10Hz (4 pulses, 1 pulse every 100ms for a total of 400 ms) with the handle pointing downwards approximately at 45° angle from the middle sagittal axis of the participants' head (Urgesi et al., 2004; Urgesi et al., 2007; Urgesi et al., 2007), adjusted to best project the pulse to the identified peak coordinate of each region and kept constant across stimulation site.

TMS targeting was managed with Brainsight 2.3.10 (Rogue Research), using individual structural and functional MRI images for each participant. The right EBA and right OPA were localized by overlaying individual activation maps from the localiser contrasts. The coil location was monitored online by the experimenter while participants performed the task, and was maintained within 1mm of the defined point. The screen displaying the participants' task was out of view of the experimenter (MG), rendering him blind to condition on a trial-by-trial basis. To ensure temporal precision, the train of TMS pulses was triggered on each trial via a TTL pulse, initiated from a photosensor which detected a screen event (unseen by participants) that co-occurred with the cue onset on each trial.

4.3.5 Stimulus creation

Pictures of bodies were obtained through internet searches and were the same used in (Johnstone and Downing, 2017). The pictures were grayscaled and cropped to exclude head and lower legs. Each body picture was presented at two different sizes (600 or 400 px height) to prevent the use of the proportion of pixels as a cue for weight judgments. Image width varied freely to maintain image ratio. Sixteen heavy and 16 slim

male and female images were collected for a total of 64 images presented at two different sizes.

Pictures of scenes were obtained through internet searches. Pictures were greyscaled and resized to 450x450 px resolution. Twenty-five garden and 25 kitchen pictures were collected and rotated by 180° for a total of 100 pictures, 50 upright and 50 inverted.

4.3.6 Stimulus presentation

All stimuli were presented centrally on a 22 inches LCD monitor set at 1920x1080 resolution and a refresh rate of 60Hz. Image presentation was controlled by PsychToolbox (Brainard, 1997) running on Octave 4 (Eaton et al., 2018) for Linux OS (Version: Xubuntu 16.04).

4.3.7 Task Procedures

In the body task, participants were asked to judge on each trial the weight (“heavy” or “slim”) of each body picture, which appeared after a verbal cue to its sex (“m” or “f”). In the scene task, participants were asked to judge the orientation of each scene (“upright” or “inverted”) after a verbal cue to its content (“kitchen” or “garden”). In 80% of trials, the cue was valid – it matched the body or scene to be judged -- and in 20% of trials it was invalid. Judgments were made by pressing one of two keys (“f” or “j”) on the keyboard. Participants were instructed to respond quickly and accurately.

Each trial was preceded by a central fixation with a random duration between 1.9 and 2.9 seconds. The written verbal cue was presented at the center of the screen for 500ms and followed by the body or scene target image, which appeared for 300 ms. Four TMS pulses at 10 Hz were applied, starting at the onset of the verbal cue and finishing 200ms before image onset. Participants performed 160 trials per stimulation site for a total of 320 trials. Trial order was counterbalanced such that the full design (combination of cue type, target type, and validity) was presented in each chunk of 20 trials. To familiarise participants with the task requirements, they first performed 48 practice trials where the verbal cue was replaced with an “x”. Stimulation site was blocked, with initial site alternated across participants (11 participants started with EBA

stimulation and 10 with OPA stimulation in both tasks). Participants were invited to take a short break every 32 trials.

4.4 Statistical Analyses

Data pre-processing and analyses were conducted using R (Version 3.5.1) packages: “dplyr” (pre-processing) “ez” (ANOVAs). Effect sizes were calculated using Jamovi (Version 0.9). Figures were generated using R package “ggplot2”.

4.4.1 Analysis

Statistical significance was tested with factorial design ANOVAs and follow-up t-tests. Significance level was set at $p = 0.05$. In accord with our instructions to participants to respond quickly and accurately, we report analyses of efficiency, computed for each condition and participant as the mean of accurate response times divided by the proportion correct. This measure assesses the effects of stimulation and validity on speed and accuracy in the aggregate. Similar, albeit sometimes weaker, patterns of results were found in analyses of the mean accurate response times and of proportion correct: Site x Task x Validity mixed-design ANOVA on accurate response times, $F(1, 40) = 2.11$, $p=0.15$; on proportion correct, $F(1,40) = 5.83$, $p=0.02$. Descriptive statistics for all measures (efficiency, accurate response time, and proportion correct) are detailed in Table S4.2.

4.5 Data and Software Availability

Raw data for this article are accessible via OSF via this link: (<https://osf.io/cysw3/>)

5. General Discussion

Over the last decade, our understanding of the visual processing of social stimuli has benefitted from the adoption of the emerging social vision approaches. The large majority of contributions using a social vision approach has focused on the perception of the face. Perhaps thanks to such contributions, to date, researchers have reached a certain degree of consensus over a standard model of face perception (Haxby et al., 2000; Duchaine and Yovel, 2015). The empirical work presented in this thesis aimed to extend social vision approaches to the case of human body perception. Taken together, the results of each chapter establish the building blocks for a perceptual model of the human body which integrates the extraction of socially relevant information from the appearance of the body. In the present chapter, I first provide a summary of the empirical findings from chapter 2, 3, and 4. I discuss how these findings improve our understanding of body perceptual representations. I then propose how they contribute towards a model of human body perception and how the logic of paradigms employed in this thesis inform the community about general mechanisms of high-level vision. Finally, I present an outline of the prospective mechanisms that account for rapid social categorisation during the perception of the social stimuli.

5.1 Summary of findings

The empirical chapters included in this thesis have investigated whether perceptual representations of the body are modulated by the socially relevant information they convey.

In chapter 2 I investigated the perceptual encoding of male and female body shapes. Specifically, by drawing a parallel with evidence for a male decision bias outlined in the social psychological literature, I tested the hypothesis that female body shapes are visually encoded as a departure from “default” male body shapes. This set of studies investigated how male and female body shapes are related to each other using a visual task. Across six experiments I demonstrate a visual search asymmetry for sex from body shape. Female bodies were more readily found among males than the converse. This pattern of results held across different views, across participants’

gender, and did not depend on low-level visual properties of the images but rather emerged from implied whole-body representations. These findings unravel the structure of body representations within the socially relevant dimension of sex and improve our understanding of how people use body cues to infer others' sex.

In chapter 3 I investigated the structure of body representations behind two highly salient socially relevant dimensions: body sex and body emotional expression. Here, I follow the logic of Garner (1974) to test the integrality or separability of processing between body sex and body emotional expression. Previous social psychological research has suggested strong associations between sex and emotion dimensions, in that emotional expression judgements are biased by sex in a stereotypical direction (e.g. female = positive emotion, Bijlstra et al., 2018). In this set of studies, I investigate whether such biases reflect deeper interactions between sex and emotion on a perceptual level. Across three experiments I used the Garner selective attention task to test whether such biased relations between body sex and body emotion judgment hold on a perceptual level. Indeed, in this task, any potential interaction that body sex exerts on body emotion processing (or vice-versa) is assessed while keeping one of the two dimensions under investigation fully irrelevant for the decision criterion. Results showed absence of Garner interference and were interpreted in favour of independent processing of body sex and body emotional expression. These findings reveal something new about the mental architecture behind the perception of some socially relevant characteristics from the body. Moreover, they draw important distinctions from previous Garner experiments which investigated facial sex and emotion (Atkinson et al., 2005; Becker, 2017).

In chapter 4, I directly test an overarching proposal of social vision approaches. That is, the notion that perceptual representations are sensitive to socially relevant information. Using TMS, I demonstrated that socially relevant knowledge about a body characteristic, elicited by forming an expectation over an incoming visual target, directly modulates the recruitment of body perceptual representations. Presenting a verbal cue indicative of the sex of an incoming picture of a body has been shown to generate validity effects in weight judgments of that image (Johnstone and Downing, 2017; Gandolfo and Downing, 2019). Delivering TMS over body selective region EBA at cue onset removed the validity effects found in this task. The importance of this finding is

twofold: On one hand, the data demonstrate that body perceptual region EBA is directly involved in the interaction between body sex and weight information and, in turn, reveals that body perceptual regions are directly involved in coding expectations about a socially relevant characteristic of the body (i.e. sex). On the other, this result directly speaks to the neural mechanisms of perceptual expectations in general. That is, in a second TMS experiment, I demonstrate that expectancy effects towards the content of a scene are disrupted following stimulation of scene selective region OPA. Under the assumption that social categorisation processes are a form of established, long-term expectations towards the visual appearance of other people (Macrae and Bodenhausen, 2000), the utility of social vision approaches for understanding general cognitive processes becomes clear. Indeed, the findings presented in this chapter contribute to a long tradition of research that has been investigating how attentional processes modulate high-level visual regions (O'Craven et al., 1999; Peelen et al., 2009; Battistoni et al., 2017; Hickey and Peelen, 2015), and to a general understanding of the properties of category selective brain regions (Peelen and Downing, 2017; Taylor and Downing, 2011; Op de Beeck et al., 2019; Battistoni and Peelen, 2016).

5.2 Towards a model of human body perception

The empirical findings I have reported improve our understanding of the cognitive architecture behind the perceptual processing of the human body.

The ultimate goal of cognitive models is to understand the unseen psychological events that occur between a signal in input (e.g. a face, or a written word) and a successful recognition and/or response to this signal (e.g. recognising a face, producing a sound). With respect to social stimuli, influential models of face perception, established on the basis of neuropsychological, cognitive and neuroimaging findings, have provided a unified framework for understanding the nature of the computations, and the brain correlates, underlying efficient recognition of faces (Bruce and Young, 1986; Young and Bruce, 2011; Haxby et al., 2000; Duchaine and Yovel, 2015).

This thesis attempts to gather previous findings of body perception, and the research reported in the empirical chapters, into a neurocognitive model deliberately inspired on existing perceptual models of face perception. There are at least two

reasons to draw parallels with existing models of face perception: Firstly, models of face perception (Bruce and Young, 1986; Young and Bruce, 2011) have successfully borrowed from theorising in other fields of psychology. For example, the familiar versus unfamiliar distinction proposed by Bruce and Young (1986) was inspired by existing dual-route box-arrow models of reading (Baron, 1977; Coltheart, 1980). Secondly, and perhaps most obviously, faces and bodies are nearly always perceived together in space and time, and carry similar information about their conspecifics (sex, age, race, etc.). Further, face and body processing show a similar, yet convincingly dissociable, neuroanatomical organisation along the posterior-lateral and ventral-medial portions of the occipito-temporal lobes (Downing et al., 2005; Peelen and Downing, 2005; Schwarzlose et al., 2005; Taylor and Downing, 2011), and a similar, yet dissociable electrophysiological timecourse (Thierry et al., 2006).

In an attempt to unify the findings from human body perception into a common framework, parallels with models of face perception have been previously made (Minnebusch and Daumn, 2009; Taylor and Downing, 2007). For example, Taylor and Downing (2007), following the logic proposed by Haxby et al. (2000) hypothesised a hierarchical scheme between lateral-occipital (EBA/OFA) and ventral-medial (FBA/FFA) face- and body-selective regions, in that EBA activity is driven by individual features of the body (body parts) and FBA activity by the visual appearance of ensembles of body parts.

In keeping with the tradition of drawing parallels with the knowledge of face perceptual processes, I provide the foundational behavioural work needed to progress towards a unified framework of human body perception. To achieve this, I follow the most recent proposals of social vision based on the face and focus on how our mental representations of the human body capture and encode body characteristics associated with socially relevant cues.

5.3 Encoding of sex from body shape

A core prediction of social vision approaches relates to high integration between socially relevant high-level constructs (e.g. stereotypes, impressions) and perceptual

representations of the stimuli on which such high-level constructs are formed (Freeman and Ambady, 2011; Freeman and Johnson, 2016; Adams et al., 2011).

Previous studies on the perception of body sex have established that several cues related to body shape (and particularly waist-to-hip ratio) support categorical judgments of sex and gender (Johnson and Tassinari, 2005). Further, visual adaptation studies provide evidence that male and female body sexes are represented primarily by two opponent pools of cells broadly tuned to the opposite sexes (Palumbo et al., 2012; Ghuman et al., 2010; Winckler and Rhodes, 2005). In agreement with the idea that after-effects found with high-level properties reflect stimulus-specific perceptual processes (Webster and Macleod, 2011), adaptation effects of sex from body parts modulate the N170 ERP component (Kovacs et al., 2005). While these studies strongly suggest the presence of distinct detectors for the two sexes, they cannot directly inform about how male and female body representations are related to each other. One possibility comes from classic interpretations of high-level adaptation effects (Leopold et al., 2001; Rhodes and Jeffery, 2006; Rhodes et al., 2005), in that male and females are organised by following a fully symmetrical opponent coding system. Accordingly, opposing after-effects for male and female would be explained by the presence of “two channels” tuned to opposite poles of the dimension of sex. In this view, the norm (in this example, the androgynous body) occurs at the unique point where the responses of the two channels are balanced, thus equidistant. Another possibility comes from proposals of social psychological literature, which instead suggest a decision bias for male bodies: unless unequivocally female, decisions about a simplified body shape will be male (Freeman et al., 2012).

Following the logic of the social vision approach, in chapter 2, I demonstrated that such decision bias is reflected in the perceptual encoding of sex from body shape. Female bodies are defined by the presence of additional, positive perceptual evidence; unless such evidence is present the encoding of body shape will be set on a male interpretation.

How does the search asymmetry inform us about the perceptual encoding of the two sexes? A possible answer to this question is provided by classic accounts of prototype-deviation asymmetries, based on pooled response models posited by

Treisman and Gormican (1988). These models generally hypothesise that detectors maximally sensitive to standard or reference values are more strongly activated by off-standard values than detectors for nonstandard values are by standard values. In this sense, female bodies activate the detectors for male, whereas male bodies produce none or little effect on the female detectors. This can be the case for two possible reasons: 1) male detectors are more broadly tuned than detectors for female bodies. 2) for asymmetric inhibition, in that male detectors, in response to their preferred stimulus, would inhibit more strongly the female detectors than the converse. Both these models predict that the deviating distractors (females) would generate high background activity in detectors for males, and deviating targets (females) would be more detectable by virtue of their additional unique activity in their own, more narrowly tuned detectors.

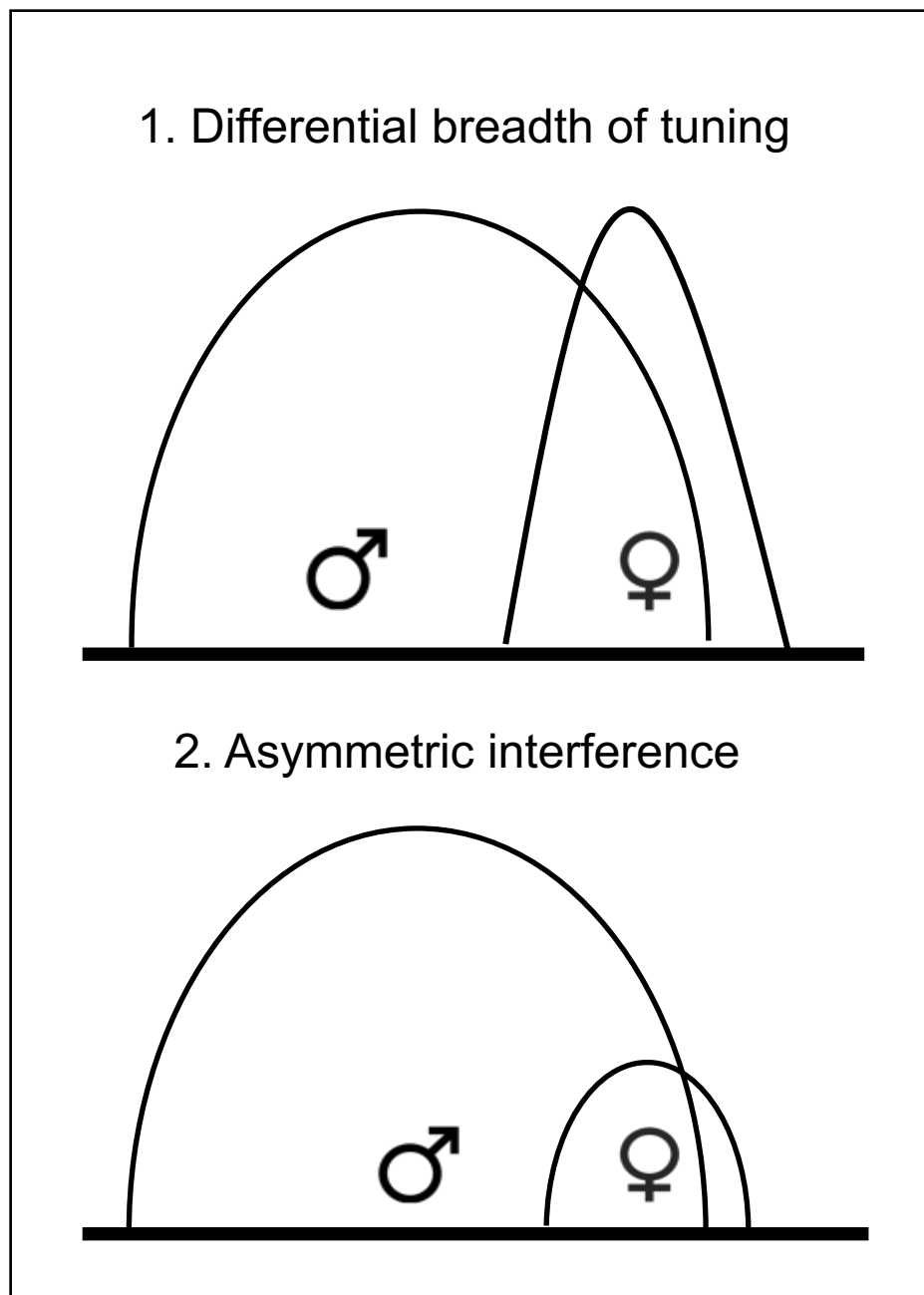


Figure 5.1 Pooled response models accounting for the search asymmetry of body sex. 1) Differential tuning breadth of male and female detectors. 2) Asymmetrical interference between male and female channels. Male channels operate greater interference on female channels than the converse. Importantly, both models have the same prediction: both models predict that female, when distractors, generate high background activity in the male channels. When targets, females are more readily found because of their additional activity in their more narrowly tuned channels.

I speculate that a differential breadth of tuning between male and female detectors is likely to be explained by an established property of male/female detectors

(in line with (1), described above and in figure 5.1) rather than being attributable to asymmetric inhibition that male exert over female detectors. A mechanism that could account for such an organisation consists of a differential expertise with female versus male caregivers during early development (Quinn et al., 2005). Indeed, an early higher exposure to female bodies (and faces) could account for preferences for female face and body stimuli found during development (Ramsey-rennels and Langlois, 2006; Alexander 2016), and for the search asymmetries I reported in adults. These preferences might be explained by narrower, and thus sharper, perceptual channels for female stimuli, formed through experience during early development. Indeed, previous studies have reported that detection performance in search tasks was related to the amount of expertise for the search target (i.e. bird, car or face) as measured by a discrimination task (Reeder et al., 2016; Herschler and Hochstein, 2009). Expertise effects have also been reported with fMRI, showing category-selective responses to objects of expertise (Gauthier et al., 2000; McGugin et al., 2012). These additional, selective representations for specific targets brought by extensive experience may account for facilitated detection processes.

However, it is also possible that the differential breadth of tuning between male and female detectors is expressed via asymmetrical inhibition that the “default” male value exerts on the “non-standard” female values (i.e. (2) outlined above, and in bottom figure 5.1). This becomes plausible if we conceive the “default” value as the most expected stimulus in the environment (even without being the most experienced stimulus during development). The sharpening hypothesis of expectations (Rumelhart and McClelland, 1982; Kok et al., 2012; Smit and Muckli, 2010) proposes that the representation of the expected sensory signals, or the “default” value, inhibits inputs that are inconsistent with the expected ones (“the non-standard value”). Accordingly, neuronal population encoding not the expected (or “default” value), but the unexpected (“non-standard”) value results in a sharper, more selective representation. A similar physiological mechanism has also been proposed to underlie feature-based attention (Martinez-Trujillo and Treue, 2004) or adaptation effects (Summerfield et al., 2008; McMahon and Olson, 2007). For example, repetition suppression has been interpreted as a reduction in prediction error. When stimulus repetitions were unlikely (and thus unexpected) repetition suppression was found to be lower (Summerfield et al., 2008). Conversely, when repetitions occurred more frequently in a block, participants showed

increased repetition suppression. The brain “predicts” the statistical regularities in the flow of incoming sensory information and has reduced processing demands for the most expected stimulus. If the encoding of male and female bodies functions in virtue of this mechanism then male bodies would be the most “expected” stimulus for which the visual system would be constantly “adapted”. When a female body is perceived a release from adaptation would occur resulting in increased/sharper activity in response to the “non-standard”, unexpected value.

5.3.1 Encoding of body sex in the brain

Although several studies have described the functional properties of body selective regions in the brain (Minnebusch and Daum, 2009; Peelen and Downing, 2007; 2011), the modulation of responses in these regions as a function of the sex of the observed bodies has not been directly assessed per se. Whenever this has been investigated, the sex of the body stimuli was not directly the factor of interest, instead these studies have investigated the effects of sexualised/non-sexualised stimuli (e.g. Cikara et al., 2011; Bernard et al., 2017; 2018), or the effect of stimuli sex with respect to that of the observer (Aleong and Paus, 2010), or of sex with respect to weight cues (Foster et al., 2019). For example, Cikara et al. (2011) report higher activation in (not functionally localised) fusiform gyri when participants were attending to female sexualised stimuli. This effect is difficult to transfer to a general difference between male/female targets because only male participants and whole-person natural images were used. Similarly, Bernard et al. (2018ab; 2019), although in this case across sexualised and non-sexualised targets, in three studies report larger N170 amplitude when participants (male and female) were attending to female versus male stimuli. The authors, however, do not interpret this main effect nor attempt to control for the male versus female interclass variability of their stimuli. This might hamper the specificity of this effect for body sex (Thierry et al., 2007).

A sharper tuning of female body perceptual channels would also be consistent with what has been indirectly reported in the findings mentioned above (Cikara et al., 2011; Bernard et al., 2018ab; 2019). Specifically, by virtue of the additional features that are engaged with reference to male bodies, female bodies will elicit increased neural activity in body selective regions. In parallel, such as hypothesised neural effect also fits

with its possible origin in early development of visual expertise (Quinn et al., 2002). Computational neurobiological approaches investigating sensory systems under the efficient coding hypothesis (Simoncelli, 2003) have examined the role of established priors over the simulated response of sensory neurons (Ganguli and Simoncelli, 2014). Specifically, the firing rate of sensory neurons in response to the most frequently occurring stimuli, was simulated theorising a higher number of cells, with a corresponding narrower tuning width. In turn, this firing rate accounted for higher perceptual sensitivity for those stimuli (Ganguli and Simoncelli, 2014). Based on this view, stimuli which have stronger priors, as female bodies, generate sharper/narrower perceptual channels and higher perceptual sensitivity.

The body selective region crucial for the asymmetrical encoding of body sex remains unknown. There are, however, at least two lines of evidence which suggest that body selective region EBA might be involved in extracting the information of body sex. One comes from evidence that social categorical information, including sex, can be rapidly extracted from simple facial (Macrae and Martin 2007; Brown and Perrett, 1993; Hill et al., 1995; Yamaguchi et al., 2013; Chronicle et al., 1995) and body cues (Johnson and Tassinary, 2005) and it is resilient to common manipulations that disrupt configural processing such as inversion and blurring (Cloutier and Macrae, 2007; Mason et al., 2006). If part-based rather than configural processing is involved in extracting sex information, then it is likely that EBA, more strongly involved in the processing of body parts (Taylor and Downing, 2007; Urgesi et al., 2007), encodes the sex from the body. This view is challenged by the findings of experiment 4 in chapter 2, where I demonstrate that the search asymmetry of sex emerges from whole body representations as it does not occur with inverted iconic stimuli. However, it is possible that when a realistic, natural stimulus is attended, people can easily extract the information of sex from a single (or a small number) of cues or body parts. The other comes from the body perception task I report in chapter 4. Online stimulation of EBA was able to disrupt the effect that a verbal cue indicative of body sex had on further weight judgments of body images. Although it is hard to make claims about the specificity of this effect for EBA versus other body selective regions using TMS, this finding directly implicates this region in the encoding of body information relevant for extracting body sex, and making use of it for efficient body processing of subsequent visual stimuli.

In sum, perceptual representations of the body are fundamentally organised with respect to the social information of sex. Visual depictions of body sex, even when iconic (experiment 3 and 4 in chapter 2), are asymmetrically encoded, in the sense that female bodies are represented with reference to a male “default”. This pattern might reflect differential tuning breadth of neuronal populations that encode for body sex in the direction of sharper/narrower detectors for female bodies compared to males. And it is likely that such organisation belongs to body selective region EBA (and maybe other regions). Together, I propose that a perceptual model of the human body should include the dimension related to body sex. Following the basic distinctions of Haxby et al. (2000), the information of sex extracted from the shape of the body is likely to be processed within regions that: 1) belong to the core system and thus perform the visual analysis of the body; 2) encode invariant aspects of the body.

A generally interesting property of socially-relevant cues, and particularly sex, is that it can be extracted also from dynamic cues conveyed by the body (Johnson and Tassinari, 2005; Kozlowski and Cutting, 1977). Of particular note, dynamic cues of body sex are sensitive to adaptation effects, indicating that male and female motion cues might be processed by distinct neuronal populations (Troje et al., 2006). Further, in this adaptation study, the authors report that psychometric functions of sex were shifted at baseline (without adaptation), showing a bias to respond male. The subjectively androgynous walker was shifted about 1SD towards male compared to the mathematical midpoint of sex-related motion. Moreover, sex-related body dynamics also influence judgments of other motion cues such as forward/backward walking direction (Brooks et al., 2008). A limitation of chapter two, and of the conclusions made here, is that the presence of a similar search asymmetry with dynamic body stimuli conveying exclusively the sex information from their motion was not tested. Based on the evidence for similar male biased decisions with ambiguous stimuli (Troje et al., 2006), it is likely that a search asymmetry for sex from body-motion would occur.

5.3.2 Asymmetry of sex across stimulus classes and future directions

The body, static and dynamic, is not the only source from which we extract information about others' sex. Indeed, there is large body of evidence that attempts to

identify what are the critical cues for detecting sex from the face (Nestor and Tarr, 2008; Burton et al., 1993; Hill et al., 1995; Macrae and Martin 2007; Yamaguchi et al., 2013). Sex can be extracted from simple cues of the face like hairstyle (Macrae and Martin, 2007; Brown and Perrett, 1993), eyebrows (Campbell et al., 1999), nose (Chronicle et al., 1995), and colour (Dupuis-Roy et al., 2009; Hill et al., 1995; Bruce and Langton 1994). Moreover, there is evidence for a male bias when judging face silhouettes (Davidenko, 2007) or computer-generated faces (Watson et al., 2016). In line with the previous evidence discussed, and with the results of chapter 2, it is reasonable to hypothesise that a similar search asymmetry would also occur with facial stimuli.

Biases in deciding the sex of a social stimulus have been consistently reported across social stimuli, in that views for an “ubiquitous” male bias have been proposed (Gaetano et al., 2016). If a search asymmetry for sex is general, and occurs across every stimulus that conveys sex information, what are the implications of conceiving this effect as grounded in stimulus specific perceptual processes? This question raises two possibilities: Firstly, in line with parallels that show similar functional organisation of face and body selective regions (Minnebusch and Daumn, 2009; Taylor and Downing, 2011), it is possible that distinct category selective regions encode sex in a similar fashion. For example, if occipital-lateral EBA asymmetrically encodes sex cues from the body then, OFA, which previously has showed similar functional properties to EBA (Taylor and Downing, 2011), would asymmetrically encode facial properties related to sex; Secondly, asymmetrical representation of sex across stimulus type might originate from a supramodal, top-down processing that is highly integrated with the incoming information of bottom-up cues relevant to sex (Freeman and Johnson, 2016).

Future studies would then be necessary to assess whether the generality of “social” biases towards male in other stimulus classes is also reflected in asymmetrical search patterns favouring female stimuli. Further, fMRI approaches could directly test the hypothesis of a differential breadth of tuning of male and female channels on a neural level. Asymmetrical patterns of search predict that female bodies generate higher background activity in detectors for males than the converse. This arrangement presents a testable prediction using classification approaches on voxel patterns in functionally localised body selective regions. For example, a classifier trained to discriminate responses to realistic versus silhouetted bodies using only female body

stimuli, should reach higher performance when tested on male body stimuli than when the same classifier is trained on male bodies and tested on female bodies. Moreover, a classifier trained on female and male bodies then tested on a separate set of body stimuli should more accurately classify females over male bodies. Based on the hypothesis outlined above, such classification pattern might be more pronounced in the category selective region EBA compared to FBA, especially in the case realistic stimuli are used. Another future direction relates to directly testing the expertise account. If the asymmetrical pattern of search originates in the high exposure to female over male faces (and bodies) early in development, then people without such early exposure should show less, or even no asymmetry of search for females. This approach would encounter all the difficulties of having access to a population with little to no exposure to females during early development.

Currently, the possibility that asymmetry of search originates from the “innate” constraints given by the nature of the perceptual channels of body selective regions cannot be ruled out. For example, Rosch-Heider (1972; 1973) showed that even the Dani of New Guinea -a culture in which no chromatic color terms were used- found easier to recognise the prototypical colours (red, green, blue and yellow) and learn categories organised around their focal points. As for colors, it is then possible that the visual system might be already geared up with a differential breadth of tuning of detectors of male and female bodies regardless of individual developmental trajectories.

In sum, further research is necessary to unravel the underlying mechanisms and the cross-stimulus (faces, biological motion) generality of the search asymmetry for sex. The empirical work reported in this thesis strongly support that body perceptual representations are sensitive to sex information. The asymmetrical encoding of this cue revealed by the search studies allows to form clear predictions on how the distinct perceptual representations of sex might be organised in the brain.

5.4 Extracting multiple social cues from the body

Extracting socially-relevant information from the appearance of the body relies on the ability to extract and differentiate between meaningful attributes conveyed by the

same perceptual object. For example, when a person sees a body, subsequently notices its attributes, such as sex, age, weight, and race. Extracting one of these attributes from the body is such an effortless process that it is reasonable to consider that each attribute is independently perceived from the other. However, there are at least two reasons to think that this might not be the case. One comes from a long tradition of high-level vision research interested in the nature of perceptual objects. This research has demonstrated that attributes (or dimensions) of an object are not always perceived in the same way. Specifically, some of these attributes are perceived as an undivided whole (integrally), while others can be decomposed from each other (separable). This fundamental distinction between integral and separable attributes can be captured using the Garner selective attention task (Garner and Felfody, 1970; Garner, 1974). This task relies on the assumption that attributes of an object are separable if, and only if, people can deploy selective attention to one of them while fully ignoring the other.

The Garner approach has been extensively used to examine and expand the knowledge of fundamental distinctions posited by influential models of face recognition: familiar versus unfamiliar (Bruce and Young, 1986); variable vs invariant attributes of faces (Haxby et al., 2000, Ganel and Goshen-Gottstein, 2002; 2004; Schweineberger and Soukoup, 1998; Schweineberger et al., 1999; Karnadewi and Lipp, 2011; Atkinson et al., 2005; Le Gal and Bruce, 2002). These studies have revealed cross-talk during the processing of single facial attributes. In doing so, they were able to challenge a number of key assumptions of influential perceptual models of the face. For example, Ganel and Goshen-Gottstein (2002; 2004) challenged the familiar versus unfamiliar face dual-route model advanced by Bruce and Young (1986). Across two studies, Ganel and Goshen-Gottstein found that information of identity was integrally processed with the information of emotional expression and that this effect was even larger with familiar faces. These data support a single-route hypothesis. Familiarity, rather than generating a dissociation, increased the cross-talk among facial attributes. Together, this suggests that visual bottom-up features that convey socially relevant information from the face might not be independently processed during perception.

The other line of evidence that posits interactions among the several social cues conveyed by social stimuli stems from social psychology (Macrae and Bodenhausen,

2000; Quadflieg and Macrae, 2010; Fiske and Neuberg, 1990). For example, a sex cue (e.g. long hair) activates the category “woman” and the corresponding associative knowledge with this category (e.g. graceful). These associations may generate stereotypes and prejudices (Amodio and Devine, 2006; Fiske and Neuberg, 1990) and also directly affect the efficiency in which we extract these cues from the appearance of other people (Johnson et al., 2012; Hugenberg, 2005; Becker et al., 2007). Recent social vision approaches have brought attention to how perceptual operations are shaped by strongly learned associations among visual categories (Quadflieg and Macrae, 2010; Adams et al., 2011; Johnson et al., 2016). In doing so, several tasks have used visual stimuli trying to capture how top-down processes directly affect the visual recognition of the social stimuli.

The core challenge of this line of research relates to finding the measures that capture how associations between categories modulate the perceptual representations of the social stimuli, rather than solely representing decision biases due to high-level, semantical association between categories. Parallel activation between categories while attending facial stimuli has been widely investigated using mouse-tracking paradigms (Freeman et al., 2008; Freeman et al., 2009; Hehman et al., 2014ab; Stoller and Freeman 2016; Freeman 2018). In this paradigm the mouse trajectory is recorded while participants are asked to perform simple categorisation of a face using the mouse. In one study, for example, participants were asked to categorise the sex of computer-generated faces of opposite sex (and irrelevantly varying in race) appearing at the two upper corners of the screen (Johnson et al., 2012). The trajectory of the mouse toward the correct sex category was more direct when the sex was “congruent” with the stereotypical race association (e.g. black-male). While this paradigm has demonstrated to be effective in capturing the subliminal co-activation of categories and the continuous nature of this process, it cannot easily reveal if categorical associations directly have an effect during the perceptual stage of processing.

The Garner selective attention task is typically interpreted to capture processing at a perceptual level (Algom and Fitousi, 2016). The response is taken at face value while the second, orthogonal dimension conveyed by the stimulus is irrelevant for the performance (and response). Thus, it is implicitly assumed that the response criterion remains invariant across control and orthogonal conditions. The Garner selective

attention task can then reliably investigate the presence or absence of an overlap in processing between the bottom-up cues signalling the social dimensions under investigation, with little or no contamination from decision processes. When the Garner approach is used to investigate social categories it can prove principles on two levels: Firstly, it can inform about the architecture of the perceptual channels that perform the visual analysis of the social stimuli, of primary interest for the establishment of perceptual models; Secondly, it can directly test the degree to which well-known top-down associations among social categories are entangled in perceptual representations.

In chapter 3 I explored the architecture of the perceptual representations of body sex and emotional postures. The relationship between these cues is noteworthy for two reasons. On one hand, a recent study found that body sex cues affect recognition of body postures in a top-down, stereotype-congruent manner (Bijlstra et al., 2018). When asked to categorise naturalistic pictures of angry males and sad females, participants were faster than the converse (sad males and angry females). However, the same interaction was not present when participants were asked to categorise more controlled, computer-generated stimuli. This finding is in keeping with the social psychological literature that demonstrates similar top-down associations when judging emotional expressions from faces (Becker et al., 2007; Hugenberg and Sczesny, 2006); On the other hand, concerning facial stimuli, there is convincing evidence that these interactions reported by the social psychological literature are also present during perceptual operations. Results from Garner selective attention tasks report integrality (Becker, 2017) or partial-integrality (Atkinson et al., 2005) of face sex and emotional expressions. The results of chapter 3 showed independent (separable) processing of body sex and emotional expression. Participants were able to selectively attend to emotional body postures while fully ignoring irrelevant variation of sex and viceversa. This pattern is in clear discontinuity with social top-down biases reported with human bodies (Bijlstra et al., 2018) and faces (Becker et al., 2007; Bijlstra et al., 2010). Further, these results also distinguish from directly comparable findings using facial stimuli and investigating sex and emotional expressions cues using the Garner task (Atkinson et al., 2005; Becker, 2017; but see LeGal and Bruce, 2002).

The divergence between stereotypical associations and the perceptual representations of social categories evidenced throughout chapter 3 has important implications for social vision approaches. Recent models under this framework were only developed considering findings from facial stimuli alone (Freeman and Ambady, 2011). According to this view, categorisation processes, and stereotypical associations start at initial perception of bottom-up features. For example, anger expressions decrease the distance between the eyebrow and the eye. In turn, this cue also signals masculinity (Le Gal and Bruce, 2002; Becker et al., 2007). When using visual tasks, the crosstalk among bottom-up features relevant for sex and emotion categorisation might account for the interactions that indicate integrality of processing of these categories. Indeed, the interaction of sex and emotional expression cues was not evident with human body stimuli. It is possible that sex and emotional expression in bodies do not share bottom-up features to the same extent as seen with faces. These results raise questions about the generalizability of findings from the face to the body. Comprehensive social vision models need to test whether social biases extend to perceptual operations across faces, bodies and biological motion.

In sum, the Garner selective attention task provides an established and solid framework to test general assumptions of social vision approaches. Specifically, for its capacity of capturing perceptual processes while keeping the decision criterion stable across condition, it can reveal whether social biases are nested in perception. Further, the use of the Garner task for the perception of the social dimensions of the human body can contribute to draw direct parallels with existing literature of face perception for an understanding of the perceptual operations involved in social categorisation across stimulus category.

A common conception of face perception is that the processing of its different components reflects functional independence which is underpinned by anatomically separable components of a distributed face processing system in the brain. The idea of separable anatomical basis for variant versus invariant information from the face comes from converging evidence of brain lesioned patients (Schweineberger et al., 1995; Campbell et al., 1986; Humphreys et al., 1993; Parry et al., 1991), monkey neurophysiology (Hasselmo et al., 1989; Desimone, 1991), and functional brain imaging (Haxby et al., 2000). For example, recordings of single-cell responses in monkey

temporal cortex suggest that neurons responsive to emotional expression were found primarily in the cortex in the superior temporal sulcus, while neurons responsive to identity were found primarily in the inferior temporal gyrus (Hasselmo et al., 1989). Furthermore, patients with prosopagnosia show dissociable impairments for facial expressions, gender and identity recognition (Humphreys et al., 1993; Parry et al., 1991). Similarly, Haxby et al.'s (2000) model brought together evidence for independent processing of variable and invariant aspects of the face processed in the pSTS and FFA, respectively.

Although these data suggest that variable and invariant aspects of the face are independent overall, findings from the Garner paradigm have revealed more subtle ways in which variable versus invariant aspects of the face interact. Previous findings from the Garner task have reported convincingly asymmetric Garner interference (Atkinson et al., 2005; Schweinberger and Soukoup, 1998; Schweinberger et al., 1999; Karnadewi and Lipp, 2011) from invariant to changeable aspects of the face (e.g. from sex to emotional expression, Atkinson et al., 2005; from identity to emotional expression, Schweinberger and Soukoup, 1998; Schweinberger et al., 1999). Invariable aspects of the face provide useful references to perform computations of more changeable aspects of the face. For instance, in spite of the gross interindividual invariance in expressing emotional expressions (Ekman and Friesen, 1971), there might be some individual idiosyncracies in the expression of particular emotions. Therefore, in principle, optimal performance in the analysis of variable aspects of the face, such as emotional expression, might benefit if the invariable aspects are taken into account.

Indeed, recent revisions of Haxby model have conceived a crosstalk between the variable and invariant routes (Duchaine and Yovel, 2015). For example, adaptation studies using fMRI demonstrated that FFA (by hypothesis, involved mainly in processing invariant aspects of the face) showed a release from adaptation after subjects perceived a change in emotional expressions and not only after changes in identity (Fox et al., 2009; Xu and Biederman, 2010). Accordingly, a neuropsychological patient with lesion in right FFA but spared pSTS and OFA was shown to be impaired in the recognition of expression from static faces (Dalrymple et al., 2011). In sum, these findings show that predictions formed initially using the Garner selective attention task in

face perception provided a valid basis to investigate separability or integrality processing of facial components in the brain.

In agreement with this view, the results outlined in chapter 3 can be considered indicative of independent, anatomically separated, processing of body sex and body emotional expressions. What are the implications of independent processing between body sex and emotional postures for a perceptual model of the human body? Sex is a relatively invariant aspect conveyed by the body. Emotional body postures, although presented statically in chapter 3, represent more changeable aspects of body information. One possibility is that the independence of sex and emotion reflects the more general distinction between body form and body motion (Giese and Poggio, 2003; Peelen and Downing, 2006; Vangdeneugden et al., 2014). Similar to models proposed in face perception (Haxby et al., 2000), the processing of body form could be associated with activity in ventral regions (EBA; FBA) while dorsal regions (e.g. the pSTS) are more engaged by dynamic displays of the body (Downing et al., 2006; Giese and Poggio, 2003; Urgesi et al., 2007). Overall, it is possible that displays of emotional postures were independently processed from body sex information in virtue of the implied body dynamics they convey.

However, the latter interpretation is challenged by the evidence that emotional information modulates body selective visual areas (Peelen et al., 2007; Atkinson et al., 2012). Numerous studies have reported that static and dynamic body emotional expressions increase activation of lateral occipital (Grosbras and Paus, 2005; Grezes et al., 2007; Kret et al., 2011) and fusiform regions (deGelder et al., 2004; Hadjikhani and deGelder, 2003; Grosbras and Paus, 2005; Kret et al., 2011). Additional studies confirm these findings by functionally localising body selective regions and by performing multivoxel pattern analyses (Atkinson et al., 2012; Peelen et al., 2007). Specifically, patterns of activity indicative of body selectivity were positively correlated with patterns reflecting the contrast between emotional versus neutral full view (Peelen et al., 2007) or point-light (Atkinson et al., 2012) videoclips in body selective regions EBA and FBA. Importantly, a similar emotional modulation was not found in face-selective regions using facial, point-light stimuli (Atkinson et al., 2012). This result hints that the emotional modulation of body selective regions might not be a general mechanism by which emotional content evokes activity in high-level visual regions. Based on these findings, it

is possible that the independence between body sex and body emotion is then related to independent processing of different neural populations within body selective regions.

However, if coexisting neural populations coding for sex and emotional cues were found in the same regions, then it would be even more likely for their related computations to interact and thus showing full, or asymmetrical Garner interference. For example, studies of face perception have suggested that FFA (mainly involved in coding invariant aspects of the face) showed a release from adaptation not only when identity but also when emotional expression was changed (Fox et al., 2009; Xu and Biederman, 2010). This functional overlap within FFA might be the neurological underpinning that accounts for asymmetrical interference between identity and emotional expressions reported in earlier studies using the Garner paradigm (Schweineberger and Soukoup, 1998; Schweinberger et al., 1999; Karnadewi and Lipp, 2011).

The independent processing between sex and emotion found using the Garner task, might be related to the choice of static stimuli. The emotional modulation of body selective regions has been demonstrated using full or point-light dynamic stimuli (Peelen et al., 2007; Atkinson et al., 2012). There is evidence that recognition accuracy of body emotion is higher in full- and point-light dynamic displays than static depictions of emotion. Further, higher intensity of the movements that convey such emotions is associated with higher recognition accuracy and higher emotional intensity ratings in all the basic emotions except for sadness (Atkinson et al., 2004). Although the participants in chapter 3 could accurately categorise the emotions they were presented, it is possible that the static stimuli used hindered the emergence of Garner interference. Interestingly, people can also successfully extract the sex information from body dynamics (Kozslowsky and Cutting, 1977; Troje et al., 2006; Johnson and Tassinari, 2005). Notwithstanding the possibility of a floor effect due to the use of static emotional expression, it is possible that an interaction between body sex and body emotion at a perceptual stage relies on shared motion features linked to successful judgments of these two categories and not on static cues.

Body sex and body emotional expression extracted from static stimuli might not rely on the extraction of overlapping (or partially-overlapping) bottom-up shape features of the body. In a previous study, Johnstone and Downing (2017) showed asymmetrical

interference from sex to weight judgments using the Garner task. This result is interpreted similarly to previous Garner studies of the face (Schweinberger and Soukoup, 1998; Schweineberger et al., 1999; Atkinson et al., 2005). Specifically, the relatively invariant information of the body (i.e. sex) provides a reference to process body weight. Sex and weight are both extracted from the shape information, and possibly through part-based processing by relying on simple shape cues (Johnson et al., 2012). Sex and emotional expression are extracted from body shape and postural information, respectively. Postural information of the body, on the other hand, might be related to configural/whole-body processing (Downing et al., 2006; Urgesi et al., 2007). This distinction of part versus whole-body processing could account alone for the independence between body sex and emotional expression and the cross-talk between body sex and body weight. However, this more general distinction does not easily explain the presence of asymmetrical patterns of interference found by Johnstone and Downing (2017).

The Garner task applied to investigate the social dimensions conveyed by the body can directly serve the purpose of a perceptual model by trying to identify the organisation of perceptual channels involved in the encoding of the attributes conveyed by the human body. On the basis of the data presented in chapter 3, I propose that body sex and static body emotional expressions are perceived independently from each other. This suggests at least some degree of neural independence between these two types of information during perception. Independence of body sex and body emotion at a perceptual stage does not exclude interactions between these two cues at a later stage of processing, indicative of social biases (Bijlstra et al., 2018).

5.5 Perceiving the body in a social world

Humans mainly rely on vision to detect, recognize and classify others along social categories (Quadflieg and Macrae, 2010). The visual cues that provide the basis for social categorisation never occur in isolation. In natural conditions, when we look at other people, the social cues relevant for multiple social categories co-occur. When attending a body, for example, a low waist-to-hip ratio triggers the category female. The activated categorical representation (female), triggered by a simple cue, guides the encoding of further category-related information (Macrae and Bodenhausen, 2000;

Olson et al., 1996). In other words, the activated categorical representation allows expectations to form about other socially-relevant information that a stimulus conveys.

To what extent are these expectancies based on perceptual operations? Social categorisation processes could be conceived as an instance of perceptual expectations operating in person perception processes (Olson et al., 1996). The formation of social categories in the first place relies on tracking the probability of the co-occurrence of visual characteristics of a stimulus that likely correspond with certain, meaningful social categorical information. In turn, our perceptual system is able to capitalise on the information about the frequency of co-occurrence of these characteristics thus forming expectations about forthcoming sensory information.

A realistic picture of a human body conveys multiple social cues at the same time. Forming a categorical expectation about a body before it appears provides the observer with a prior probability that 1) is continuously compared with the subsequent observed input; 2) guides/biases the attention towards specific attributes of the body relevant for that category; 3) guides the processing of subsequently encountered attributes that are related to that category, but might be shared with other social categories. For example, body weight judgments (i.e. heavy or slim) are influenced by perceptual expectations instantiated by a symbolic cue indicating the sex (Johnstone and Downing, 2017; Gandolfo and Downing, 2019). That is, participants' categorical expectation towards sex facilitates judgments of another shape-related social cue extracted from the body. Heavy or slim judgments may be calibrated based on the expectation provided by the category "woman".

Which level of representation is enhanced through social-categorical expectations? Expectancy effects can result in asymmetrical patterns. For example, forming an expectation about the weight does not facilitate sex judgments (Johnstone and Downing, 2017). This raises two interesting, and related, possibilities: 1) long term expectations that originally formed the "weight" category are not highly reliable for sex judgments 2) Expectations, at least in the case of body sex and weight, mirror interactions happening on a perceptual level. Indeed, Garner interference between body sex and body weight reveal similar asymmetry. Only the least changeable aspect of the body (i.e. sex) interferes with judgments of weight (more changeable) but not the

converse (Johnstone and Downing, 2017). This relationship is plausible if we think about the nature of the information that these cues provide. Body weight is a more changeable aspect compared to body sex and thus provides a less stable basis to perform further computations. Expectancies generated by categorical knowledge, following the example above, could then be “constrained” by the architecture of body perceptual representations. In turn, it is possible that the structure of the perceptual representations of the body is formed upon the reliability of the co-occurrence of the visual cues relevant for detecting a certain category. In other words, categorical expectations represent adapted configurations of the visual characteristics that are able to shape perceptual representations and modulate further category-related information.

Conceiving social categories as a set of perceptual expectancies available to the observer reconciles proposals of social vision with a burgeoning literature in vision science which is set to understand vision in its context. Specifically, it has been investigating how low- and high- level perceptual representations are modulated by: expectations (Summerfield and de Lange, 2014; Battistoni et al., 2016; Kok et al., 2013; de Lange et al., 2018); reward (Hickey and Peelen, 2015; Hickey et al., 2015; Hickey and Van Zoest, 2012); positional regularities of an object (or multiple objects) in the scene (Kaiser et al., 2014; Kaiser et al., 2019), or multi-persons displays interacting (Walbrin and Koldewyn, 2019; Papeo et al., 2017). This broad range of research shares the proposal that regions performing the visual analysis of the stimuli in the striate and extrastriate cortex are highly sensitive to these numerous modulating factors operating during naturalistic vision.

Social categorisation processes might operate on a perceptual level as an instance of these general factors that shape responses in visual cortex under naturalistic conditions. In chapter 4, I demonstrated that expectations formed towards the social category of sex evoke activity in the visual cortex relevant for body perception, even in absence of visual input. This demonstrates that expectations formed towards socially-relevant information of a body stimulus are able to trigger the formation of perceptual templates that guide subsequent body processing. I propose that expectation-based biasing of pre-stimulus activity in body selective regions might be one mechanism accounting for rapid, effortless, social categorisation from the human body.

In a second experiment, I directly demonstrate that this effect is not specific to symbolic cues reflecting social information, but rather signifies a mechanism in which expectations operate on high-level perceptual processes. Expectations formed about the category of a scene (i.e. Kitchen; Garden) evoke activity in the visual cortex relevant for scene judgments before image onset. This demonstrates that expectations are expressed via pre-stimulus activation of extrastriate cortex. Further, expectations bias pre-stimulus signals maximally at the sensory neurons that encode the expected stimulus.

Together, the findings reported in chapter 4 strongly support the validity of social vision approaches. They suggest that a comprehensive perceptual model of the human body should consider investigating, and including, social categories as part of its “core system” because (at least some of them) directly modulate the visual analysis of the body. Further, these results demonstrate that social vision approaches can contribute to shed a light on the general overarching questions about vision and its function. Classical theories of object recognition would typically assume that the ultimate goal of high-level vision is to form invariant object representations (Marr, 1982). However, the large amount of information that humans are exposed to (for example, in a social situation, in front of a real person) result in an high load for our sensory and cognitive system. Mechanisms like perceptual expectations, equip the observer with highly adapted perceptual templates which reduce neural competition and allow to make optimal use of limited cortical resources. This aids the person perception process in the social domain, but also modulates, more generally, visual processes in their real-world context.

5.6 Concluding remarks

The human body conveys numerous social cues. This thesis aimed to establish the building blocks of a comprehensive model of human body perception. I proposed that this can be achieved by including the study of the perceptual operations involved in extracting social-categorical information from the human body. Further, in line with social vision proposals, the work presented and reviewed here brings together findings strictly related to the visual processing of the body, with findings more focused on the social information it conveys. These two aspects have been studied in relative isolation.

However, for a consensus model of the perception of the human body, a more general framework that integrates the “visual” and the “social” of the body is necessary.

I propose that the research approach used here successfully endorses the social vision program. It brings together the two disciplines by 1) testing interactions among social categories emerged from social psychological perspectives using established and well-understood visual tasks; 2) framing rapid social categorisation processes as an instance of highly stable visual priors formed over socially-relevant visual cues. These priors generate visual templates that bias neural activity and guide further visual processing. This functioning in the social domain resemble mechanisms operating during naturalistic vision in general.

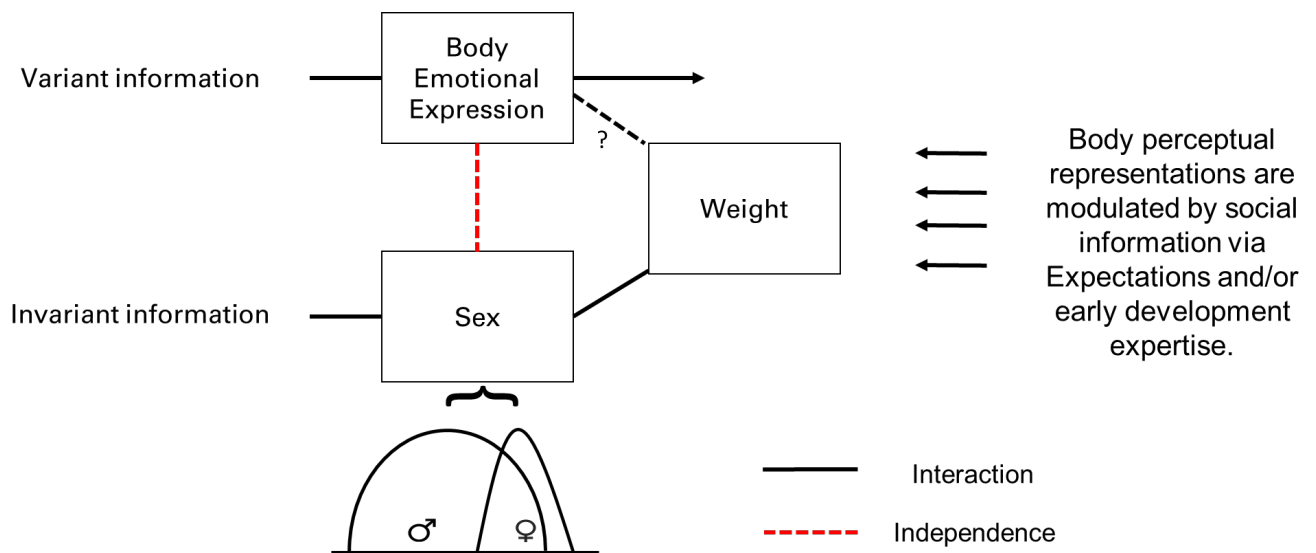


Figure 5.2. Shows an initial proposal of a social vision model of the human body.

In figure 5.2, based on the findings reported in chapter 2, 3, and 4, I attempt to draw the initial blocks of a social vision model of the body. This model does not directly report which body selective regions encode the body information under investigation. However, importantly, it directly includes the social categorical information conveyed by the body as part of a “core system” for body processing. The “core system” comprises regions directly involved in the visual processing of the body. Further, although not comprehensive of all the social categories the body can convey, this model draws a picture of the main interactions among several social cues brought by the body. The model posits differential tuning breadth of male and female body detectors within the

dimension of sex. Further, it posits independence of processing between the dimension of body sex and the dimension of body emotional expression. Moreover, it advances a one-sided interaction from the dimension of body sex to the dimension of body weight. Overall, the perceptual representations of the body that convey social categorical information appear to be shaped by general factors that modulate high level, category specific perceptual representations, such as expectations and/or early development expertise.

While the present work has focused on social categories, recent evidence suggests that humans can also extract traits from body features alone (Hu et al., 2018), and not only from the face (Todorov et al., 2015). A comprehensive social vision model of the human body will need to theorise about the perceptual operations involved in the extraction of any information that can be extracted from the body, including personality traits. Further, the motion signals conveyed by the body also carry a rich amount of social information. Here, the primary focus has been mainly on the static aspects of the body. A comprehensive model will need to include the dynamic aspects of the human body, and the social information extracted from them.

The human body is, obviously, not the only stimulus that conveys social information. The present findings may contribute more broadly to the goal of forming general models of person perception. Ideally, such a common framework will include the extraction of the full range of social signals (traits, emotions, and social categories) from the whole range of stimuli that conveys them: bodies and their motion, faces, and voices (Yovel and Belin, 2013; Freeman, Stolier and Brooks, *in press*).

References

- Adams Jr, R. B., Adams, R. B., Ambady, N., Shimojo, S., & Nakayama, K. (Eds.). (2011). *The Science of Social Vision: The Science of Social Vision* (Vol. 7). Oxford University Press.
- Adams, R. B., & Kveraga, K. (2015). Social vision: functional forecasting and the integration of compound social cues. *Review of philosophy and psychology*, 6(4), 591-610.
- Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, 393(6684), 470.
- Alexander, G. M., Hawkins, L. B., Wilcox, T., & Hirshkowitz, A. (2016). Infants prefer female body phenotypes; Infant girls prefer they have an hourglass shape. *Frontiers in psychology*, 7, 804.
- Algom, D., & Fitousi, D. (2016). Half a century of research on Garner interference and the separability–integrality distinction. *Psychological Bulletin*, 142(12), 1352.
- Allman, J. M., & Kaas, J. H. (1974). The organization of the second visual area (V II) in the owl monkey: a second order transformation of the visual hemifield. *Brain research*, 76(2), 247-265.
- Allport, G. W. (1954). (1979). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. In *Advances in experimental social psychology* (Vol. 32, pp. 201-271). Academic Press.
- Ambrus, G. G., Amado, C., Krohn, L. and Kovács, G. (2019). TMS of the occipital face area modulates cross-domain identity priming. *Brain Struct. Funct.* 224, 149-157.
- Amishav, R., & Kimchi, R. (2010). Perceptual integrality of componential and configural information in faces. *Psychonomic Bulletin & Review*, 17(5), 743-748.
- Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: evidence for independent constructs and unique effects on behavior. *Journal of personality and social psychology*, 91(4), 652.
- Andrews, T. J., Watson, D. M., Rice, G. E., & Hartley, T. (2015). Low-level properties of natural images predict topographic patterns of neural response in the ventral visual pathway. *Journal of Vision*, 15(7), 3-3.
- Anzellotti, S., & Caramazza, A. (2017). Multimodal representations of person identity individuated with fMRI. *Cortex*, 89, 85-97.

Arizpe, J. M., McKean, D. L., Tsao, J. W., & Chan, A. W. Y. (2017). Where you look matters for body perception: Preferred gaze location contributes to the body inversion effect. *PloS one*, 12(1), e0169148.

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, 41(3), 258.

Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature neuroscience*, 7(5), 542.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6), 717-746.

Atkinson, A. P., Heberlein, A. S., & Adolphs, R. (2011). Are people special? A brain's eye view. *The science of social vision*, 363-392.

Atkinson, A. P., Tipples, J., Burt, D. M., & Young, A. W. (2005). Asymmetric interference between sex and emotion in face perception. *Perception & Psychophysics*, 67(7), 1199-1213.

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(2), 1700-1712.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, 338(6111), 1225-1229.

Axelrod, V., & Yovel, G. (2015). Successful decoding of famous faces in the fusiform face area. *PloS one*, 10(2), e0117126.

Axelrod, V., & Yovel, G. (2015). Successful decoding of famous faces in the fusiform face area. *PloS one*, 10(2), e0117126.

Bagnis, A., Celeghin, A., Mosso, C. O., & Tamietto, M. (2019). Toward an integrative science of social vision in intergroup bias. *Neuroscience & Biobehavioral Reviews*.

Bannerman, R. L., Milders, M., De Gelder, B., & Sahraie, A. (2009). Orienting to threat: faster localization of fearful facial expressions and body postures revealed by saccadic eye movements. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb-2008.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American psychologist*, 54(7), 462.

Baron, J. (1977). Mechanisms for pronouncing printed words: Use and acquisition. *Basic processes in reading: Perception and comprehension*, 175-216.

Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, 58(1), 71-78.

Battistoni, E., Stein, T. and Peelen, M. V. (2016). Preparatory attention in visual cortex. *Ann. N. Y. Acad. Sci.* 1396, 92-107.

Baudouin, J. Y., & Humphreys, G. W. (2006). Configural information in gender categorisation. *Perception*, 35(4), 531-540.

Bauser, D. A. S., Suchan, B., & Daum, I. (2011). Differences between perception of human faces and body shapes: Evidence from the composite illusion. *Vision research*, 51(1), 195-202.

Bayet, L., Pascalis, O., Quinn, P. C., Lee, K., Gentaz, É., & Tanaka, J. W. (2015). Angry facial expressions bias gender categorization in children and adults: behavioral and computational evidence. *Frontiers in psychology*, 6, 346.

Becker, D. V. (2017). Facial gender interferes with decisions about facial expressions of anger and happiness. *Journal of Experimental Psychology: General*, 146(4), 457.

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: happy faces, not angry faces, are more efficiently detected in single-and multiple-target visual search tasks. *Journal of Experimental Psychology: General*, 140(4), 637.

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of personality and social psychology*, 92(2), 179.

Bergmann, T. O. (2018). Brain State-Dependent Brain Stimulation. *Front. Psychol.* 9, 2108.

Berlucchi, G., & Aglioti, S. M. (2010). The body in the brain revisited. *Experimental brain research*, 200(1), 25.

Bernard, P., Content, J., Deltenre, P., & Colin, C. (2018). When the body becomes no more than the sum of its parts: the neural correlates of scrambled versus intact sexualized bodies. *Neuroreport*, 29(1), 48-53.

Bernard, P., Hanoteau, F., Gervais, S., Servais, L., Bertolone, I., Deltenre, P., & Colin, C. (2019). Revealing clothing does not make the object: ERP evidences that cognitive objectification is driven by posture suggestiveness, not by revealing clothing. *Personality and Social Psychology Bulletin*, 45(1), 16-36.

Bernard, P., Rizzo, T., Hoonhorst, I., Deliëns, G., Gervais, S. J., Eberlen, J., ... & Klein, O. (2018) B. The neural correlates of cognitive objectification: An ERP study on the body inversion effect associated with sexualized bodies. *Social Psychological and Personality Science*, 9(5), 550-559.

Bernard, Philippe, Tiziana Rizzo, Ingrid Hoonhorst, Gaétane Deliëns, Sarah J. Gervais, Julia Eberlen, Clémence Bayard, Paul Deltenre, Cécile Colin, and Olivier Klein. (2018) A. "The neural correlates of cognitive objectification: An ERP study on the body inversion effect associated with sexualized bodies." *Social Psychological and Personality Science* 9, no. 5: 550-559.

Bestelmeyer, P. E., Jones, B. C., DeBruine, L. M., Little, A. C., & Welling, L. L. (2010). Face aftereffects suggest interdependent processing of expression and sex and of expression and race. *Visual Cognition*, 18(2), 255-274.

Bhatt, R. S., Hock, A., White, H., Jubran, R., & Galati, A. (2016). The development of body structure knowledge in infancy. *Child development perspectives*, 10(1), 45-52.

Bi, Y., Wang, X., & Caramazza, A. (2016). Object domain and modality in the ventral visual pathway. *Trends in cognitive sciences*, 20(4), 282-290.

Biederman, I., Mezzanotte, R. J. and Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cogn. Psychol.* 14, 143-177.

Bijlstra, G., Holland, R. W., & Wigboldus, D. H. (2010). The social face of emotion recognition: Evaluations versus stereotypes. *Journal of Experimental Social Psychology*, 46(4), 657-663.

Bijlstra, G., Holland, R. W., Dotsch, R., & Wigboldus, D. H. (2018). Stereotypes and prejudice affect the recognition of emotional body postures. *Emotion*.

Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, A. M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 36(6), 1477.

Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of personality and social psychology*, 55(5), 726.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. *Stereotype activation and inhibition: Advances in social cognition*, 11, 1-52.

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience*, 16(11), 3737-3744.

- Bracci, S., & de Beeck, H. O. (2016). Dissociations and associations between shape and category representations in the two visual pathways. *Journal of Neuroscience*, 36(2), 432-444.
- Bracci, S., Ritchie, J. B., Kalfas, I., & de Beeck, H. O. (2019). The ventral visual pathway represents animal appearance over animacy, unlike human behavior and deep neural networks. *Journal of Neuroscience*, 1714-18.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, 10(4), 433-436.
- Brandman, T., & Yovel, G. (2010). The body inversion effect is mediated by face-selective, not body-selective, mechanisms. *Journal of Neuroscience*, 30(31), 10534-10540.
- Brandman, T., & Yovel, G. (2012). A face inversion effect without a face. *Cognition*, 125(3), 365-372.
- Brebner, J. L., Martin, D., & Macrae, C. N. (2009). Dude looks like a lady: Exploring the malleability of person categorization. *European Journal of Social Psychology*, 39(6), 1109-1119.
- Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and social psychology bulletin*, 17(5), 475-482.
- Brooks, A., Schouten, B., Troje, N. F., Verfaillie, K., Blanke, O., & van der Zwan, R. (2008). Correlated changes in perceptions of the gender and orientation of ambiguous biological motion figures. *Current Biology*, 18(17), R728-R729.
- Brooks, K. R., Clifford, C. W., Stevenson, R. J., Mond, J., & Stephen, I. D. (2018). The high-level basis of body adaptation. *Royal Society open science*, 5(6), 172103.
- Brooks, K. R., Mond, J. M., Stevenson, R. J., & Stephen, I. D. (2016). Body image distortion and exposure to extreme body types: contingent adaptation and cross adaptation for self and other. *Frontiers in neuroscience*, 10, 334.
- Brooks, K. R., Mond, J., Mitchison, D., Stevenson, R. J., Challinor, K. L., & Stephen, I. D. (2019). Looking at the Figures: Visual Adaptation as a Mechanism for Body-Size and-Shape Misperception. *Perspectives on Psychological Science*, 1745691619869331.
- Brosch, T., Bar-David, E., & Phelps, E. A. (2013). Implicit race bias decreases the similarity of neural representations of black and white faces. *Psychological science*, 24(2), 160-166.
- Brown, E., & Perrett, D. I. (1993). What gives a face its gender?. *Perception*, 22(7), 829-840.
- Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception*, 23(7), 803-822.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British journal of psychology*, 77(3), 305-327.

Bruce, V., & Young, A. (2013). *Face perception*. Psychology Press.

Bruce, V., Burton, A. M., Hanna, E., Healey, P., Mason, O., Coombes, A., ... & Linney, A. (1993). Sex discrimination: how do we tell the difference between male and female faces?. *perception*, 22(2), 131-152.

Bruce, V., Burton, A. M., Hanna, E., Healey, P., Mason, O., Coombes, A., ... & Linney, A. (1993). Sex discrimination: how do we tell the difference between male and female faces?. *perception*, 22(2), 131-152.

Bruner, J. S. (1957). On perceptual readiness. *Psychological review*, 64(2), 123.

Bruner, J. S., & Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of personality*, 18(2), 206-223.

Bruner, J. S., & Tagiuri, R. (1954). *The perception of people*. HARVARD UNIV CAMBRIDGE MA LAB OF SOCIAL RELATIONS.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.

Burton, A. M., Bruce, V., & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception*, 22(2), 153-176.

Busigny, T., Van Belle, G., Jemel, B., Hosein, A., Joubert, S., & Rossion, B. (2014). Face-specific impairment in holistic perception following focal lesion of the right anterior temporal lobe. *Neuropsychologia*, 56, 312-333.

Buxbaum, L. J., & Branch Coslett, H. (2001). Specialised structural descriptions for human body parts: Evidence from autotopagnosia. *Cognitive Neuropsychology*, 18(4), 289-306.

Caldara, R., Seghier, M. L., Rossion, B., Lazeyras, F., Michel, C., & Hauert, C. A. (2006). The fusiform face area is tuned for curvilinear patterns with more high-contrasted elements in the upper part. *Neuroimage*, 31(1), 313-319.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, 6(8), 641.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience*, 2(5), 352.

Campbell, R, Benson, P J, Wallace, S B, Doesbergh, S, Coleman, M, 1999 “More about brows: How poses that change brow position affect perceptions of gender” *Perception* 28 489–504

Campbell, R., LANDIS, T., & REGARD, M. (1986). Face recognition and lipreading: A neurological dissociation. *Brain*, 109(3), 509-521.

Candidi, M., Urgesi, C., Ionta, S., & Aglioti, S. M. (2008). Virtual lesion of ventral premotor cortex impairs visual perception of biomechanically possible but not impossible actions. *Social neuroscience*, 3(3-4), 388-400.

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Res.* 51, 1484-1525.

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

Cazzato, V., Mian, E., Serino, A., Mele, S., & Urgesi, C. (2015). Distinct contributions of extrastriate body area and temporoparietal junction in perceiving one’s own and others’ body. *Cognitive, Affective, & Behavioral Neuroscience*, 15(1), 211-228.

Challinor, K. L., Mond, J., Stephen, I. D., Mitchison, D., Stevenson, R. J., Hay, P., & Brooks, K. R. (2017). Body size and shape misperception and visual adaptation: An overview of an emerging research paradigm. *Journal of International Medical Research*, 45(6), 2001-2008.

Chan, A. W., & Downing, P. E. (2011). Faces and eyes in human lateral prefrontal cortex. *Frontiers in human neuroscience*, 5, 51.

Chan, A. W., Kravitz, D. J., Truong, S., Arizpe, J., & Baker, C. I. (2010). Cortical representations of bodies and faces are strongest in commonly experienced configurations. *Nature neuroscience*, 13(4), 417.

Chan, A. W., Peelen, M. V., & Downing, P. E. (2004). The effect of viewpoint on body representation in the extrastriate body area. *Neuroreport*, 15(15), 2407-2410.

Chang, L., & Tsao, D. Y. (2017). The code for facial identity in the primate brain. *Cell*, 169(6), 1013-1028.

Chronicle, E. P., Chan, M. Y., Hawkins, C., Mason, K., Smethurst, K., Stallybrass, K., ... & Wright, K. (1995). You can tell by the nose—Judging sex from an isolated facial feature. *Perception*, 24(8), 969-973.

Chun, M. M., Golomb, J. D. and Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Ann. Rev. Psychol.* 62, 73-101.

Cikara, M., Eberhardt, J. L., & Fiske, S. T. (2011). From agents to objects: Sexist attitudes and neural responses to sexualized targets. *Journal of Cognitive Neuroscience*, 23(3), 540-551.

Clark, V. P., Parasuraman, R., Keil, K., Kulansky, R., Fannon, S., Maisog, J. M., ... & Haxby, J. V. (1997). Selective attention to face identity and color studied with fMRI. *Human brain mapping*, 5(4), 293-297.

Cloutier, J., & Macrae, C. N. (2007). Who or what are you?: Facial orientation and person construal. *European Journal of Social Psychology*, 37(6), 1298-1309.

Cloutier, J., Mason, M. F., & Macrae, C. N. (2005). The perceptual determinants of person construal: Reopening the social-cognitive toolbox. *Journal of Personality and Social Psychology*, 88(6), 885.

Cloutier, J., Turk, D. J., & Neil Macrae, C. (2008). Extracting variant and invariant information from faces: The neural substrates of gaze detection and sex categorization. *Social Neuroscience*, 3(1), 69-78.

Cohen, M. A., Alvarez, G. A., Nakayama, K., & Konkle, T. (2016). Visual search for object categories is predicted by the representational architecture of high-level visual cortex. *Journal of neurophysiology*, 117(1), 388-402.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception*, 29(8), 893-909.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & psychophysics*, 27(3), 183-228.

Condry, J., & Condry, S. (1976). Sex differences: A study of the eye of the beholder. *Child development*, 812-819.

Contreras, J. M., Banaji, M. R., & Mitchell, J. P. (2013). Multivoxel patterns in fusiform face area differentiate faces by sex and race. *PLoS One*, 8(7), e69684.

Corradi-Dell'Acqua, C., & Rumiati, R. I. (2007). What the brain knows about the body: evidence for dissociable representations. *Brain development in learning environments: embodied and perceptual advancements*. Cambridge Scholar Publishing, Newcastle, 50-64.

Cronbach, L. J. (1958). Proposals leading to analytic treatment of social perception scores. *Person perception and interpersonal behavior*, 353, 379.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281.

Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility: evaluative processing goals shape amygdala activity. *Psychological Science*, 19(2), 152-160.

Dalrymple, K. A., Oruc, I., Duchaine, B., Pancaroglu, R., Fox, C. J., Iaria, G., ... &

Barton, J. J. (2011). The anatomic basis of the right face-selective N170 IN acquired prosopagnosia: a combined ERP/fMRI study. *Neuropsychologia*, 49(9), 2553-2563.

Davenport, J. L. and Potter, M. C. (2004). Scene consistency in object and background perception. *Psychol. Sci.* 15, 559-564.

Davidenko, N. (2007). Silhouetted face profiles: A new methodology for face perception research. *Journal of Vision*, 7(4), 6-6.

Davidenko, N., Remus, D. A., & Grill-Spector, K. (2012). Face-likeness and image variability drive responses in human face-selective ventral regions. *Human brain mapping*, 33(10), 2334-2349.

de Beeck, H. P. O., Pillet, I., & Ritchie, J. B. (2019). Factors Determining Where Category-Selective Areas Emerge in Visual Cortex. *Trends in cognitive sciences*.

de Beeck, H. P. O., Pillet, I., & Ritchie, J. B. (2019). Factors Determining Where Category-Selective Areas Emerge in Visual Cortex. *Trends in cognitive sciences*.

de Beeck, H. P. O., Torfs, K., & Wagemans, J. (2008). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *Journal of Neuroscience*, 28(40), 10111-10123.

de C. Hamilton, A. F., & Grafton, S. T. (2007). Action outcomes are represented in human inferior frontoparietal cortex. *Cerebral Cortex*, 18(5), 1160-1168.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242.

De Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3475-3484.

De Gelder, B. (2016). *Emotions and the Body*. Oxford University Press.

de Gelder, B., Van den Stock, J., Meeren, H. K., Sinke, C. B., Kret, M. E., & Tamietto, M. (2010). Standing up for the body. Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neuroscience & Biobehavioral Reviews*, 34(4), 513-527.

de Lange, F. P., Heilbron, M. and Kok, P. (2018). How do expectations shape perception?. *Trends Cogn. Sci.* 22, 764-779.

De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, 13(4), 247-268.

De Renzi, E., & Scotti, G. (1970). Autotopagnosia: Fiction or reality?: report of a case. *Archives of Neurology*, 23(3), 221-227.

Decety, J., & Lamm, C. (2006). Human empathy through the lens of social neuroscience. *The scientific World journal*, 6, 1146-1163.

DeGutis, J., Chatterjee, G., Mercado, R. J., & Nakayama, K. (2012). Face gender recognition in developmental prosopagnosia: evidence for holistic processing and use of configural information. *Visual Cognition*, 20(10), 1242-1253.

Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of cognitive neuroscience*, 3(1), 1-8.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, 4(8), 2051-2062.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*, 56(1), 5.

Devue, C., Collette, F., Balteau, E., Degueldre, C., Luxen, A., Maquet, P., & Brédart, S. (2007). Here I am: the cortical correlates of visual self-recognition. *Brain research*, 1143, 169-182.

Dilks, D. D., Julian, J. B., Paunov, A. M. and Kanwisher, N. (2013). The occipital place area is causally and selectively involved in scene perception. *J. Neurosci.* 33, 1331-1336.

Dillon, M. R., Persichetti, A. S., Spelke, E. S. and Dilks, D. D. (2018). Places in the brain: Bridging layout and object geometry in scene-selective cortex. *Cereb. Cortex* 28, 2365-2374.

Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25(6), 727-738.

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology*, 22(1), 22-37.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of experimental social psychology*, 33(5), 510-540.

Downing, P. E., & Peelen, M. V. (2011). How might occipitotemporal body-selective regions interact with other brain areas to support person perception?. *Cognitive neuroscience*, 2(3-4), 216-226.

Downing, P. E., & Peelen, M. V. (2016). Body selectivity in occipitotemporal cortex: Causal evidence. *Neuropsychologia*, 83, 138-148.

- Downing, P. E., Bray, D., Rogers, J., & Childs, C. (2004). Bodies capture attention when nothing is expected. *Cognition*, 93(1), B27-B38.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Downing, P. E., Peelen, M. V., Wiggett, A. J., & Tew, B. D. (2006). The role of the extrastriate body area in action perception. *Social Neuroscience*, 1(1), 52-62.
- Downing, P. E., Wiggett, A. J., & Peelen, M. V. (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *Journal of Neuroscience*, 27(1), 226-233.
- Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, 1, 393-416.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological review*, 96(3), 433.
- Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of vision*, 9(2), 10-10.
- Eaton, J. W., Bateman, D., Hauberg, S. and Wehbring, R. (2018). GNU Octave version 4.4.1 manual. <https://octave.org/doc/v4.4.1/>
- Edwards, G. J., Taylor, C. J., & Cootes, T. F. (1998, April). Interpreting face images using active appearance models. In *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 300-305). IEEE.
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124.
- Ellis, H. D., & Florence, M. (1990). Bodamer's (1947) paper on prosopagnosia. *Cognitive neuropsychology*, 7(2), 81-105.
- Enns, J. T., & Rensink, R. A. (1990). Sensitivity to three-dimensional orientation in visual search. *Psychological Science*, 1(5), 323-326.
- Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, 98(3), 335.
- Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, 12(10), 388-396.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598.

Esterman, M. and Yantis, S. (2010). Perceptual expectation evokes category-selective cortical activity. *Cereb. Cortex* 20, 1245-1253.

Fechner, G. T. (1948). Elements of psychophysics, 1860. *Dennis, W., éditeur: Readings in the history of psychology*, 206-213.

Fisher, C., & Freiwald, W. A. (2015). Whole-agent selectivity within the macaque face-processing system. *Proceedings of the National Academy of Sciences*, 112(47), 14717-14722.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In *Advances in experimental social psychology* (Vol. 23, pp. 1-74). Academic Press.

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*. Sage.
Fiske, S. T., Gilbert, D. T., & Lindzey, G. (Eds.). (2010). *Handbook of social psychology* (Vol. 2). John Wiley & Sons.

Foster, C., Zhao, M., Romero, J., Black, M. J., Mohler, B. J., Bartels, A., & Bühlhoff, I. (2019). Decoding subcategories of human bodies from both body-and face-responsive cortical regions. *Neuroimage*, 202, 116085.

Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. (2009). The correlates of subjective perception of identity and expression in the face network: an fMRI adaptation study. *Neuroimage*, 44(2), 569-580.

Freeman, J. B. (2018). Doing psychological science by hand. *Current Directions in Psychological Science*, 27(5), 315-323.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, 118(2), 247.

Freeman, J. B., & Johnson, K. L. (2016). More than meets the eye: Split-second social perception. *Trends in Cognitive Sciences*, 20(5), 362-374.

Freeman, J. B., Rule, N. O., Adams Jr, R. B., & Ambady, N. (2009). The neural basis of categorical face perception: Graded representations of face gender in fusiform and orbitofrontal cortices. *Cerebral Cortex*, 20(6), 1314-1322.

Freeman, J. B., Stoller, R. M., & Brooks, J. A. (2019). Dynamic interactive theory as a domain-general account of social perception. (in press)

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

Freiwald, W. A., Tsao, D. Y., & Livingstone, M. S. (2009). A face feature space in the macaque temporal lobe. *Nature neuroscience*, 12(9), 1187.

Friedman-Hill, S., & Wolfe, J. M. (1995). Second-order parallel processing: Visual search for the odd item in a subset. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3), 531.

Gaetano, J., Brooks, A., & van der Zwan, R. (2012). Discriminating sex from representations of the human hand: Evidence of a pan-stimulus male bias. *Front Hum Neurosci*.

Gaetano, J., van der Zwan, R., Blair, D., & Brooks, A. (2014). Hands as sex cues: Sensitivity measures, male bias measures, and implications for sex perception mechanisms. *PloS one*, 9(3), e91032.

Gaetano, J., Van Der Zwan, R., Oxner, M., Hayward, W. G., Doring, N., Blair, D., & Brooks, A. (2016). Converging evidence of ubiquitous male bias in human sex perception. *PloS one*, 11(2), e0148623.

Gandolfo, M., & Downing, P. E. (2019). Causal evidence for expression of perceptual expectations in category-selective extrastriate regions. *Current Biology*, 29(15), 2496-2500.

Ganel, T. (2011). Revisiting the relationship between the processing of gaze direction and the processing of facial expression. *Journal of Experimental Psychology: Human Perception and Performance*, 37(1), 48.

Ganel, T., & Goshen-Gottstein, Y. (2002). Perceptual integrality of sex and identity of faces: Further evidence for the single-route hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, 28(4), 854.

Ganel, T., Goshen-Gottstein, Y., & Goodale, M. A. (2005). Interactions between the processing of gaze direction and facial expression. *Vision research*, 45(9), 1191-1200.

Ganel, T., Valyear, K. F., Goshen-Gottstein, Y., & Goodale, M. A. (2005). The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia*, 43(11), 1645-1654.

Ganguli, D., & Simoncelli, E. P. (2014). Efficient sensory encoding and Bayesian inference with heterogeneous neural populations. *Neural computation*, 26(10), 2103-2134.

Garner, W. R. (1974). The stimulus in information processing. In *Sensation and measurement* (pp. 77-90). Springer, Dordrecht.

Garner, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive psychology*, 8(1), 98-123.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, 1(3), 225-241.

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. *Journal of cognitive neuroscience*, 12(3), 495-504.

Ghuman, A. S., McDaniel, J. R., & Martin, A. (2010). Face adaptation without a face. *Current Biology*, 20(1), 32-36.

Gibson, J. J. (1966). The senses considered as perceptual systems.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston, MA, US.

Giese, M. A., & Poggio, T. (2003). Cognitive neuroscience: neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, 4(3), 179.

Gilbert, D. T. (1998). Ordinary personology. *The handbook of social psychology*, 2, 89-150.

Gilboa-Schechtman, E., Ben-Artzi, E., Jeczemien, P., Marom, S., & Hermesh, H. (2004). Depression impairs the ability to ignore the emotional aspects of facial expressions: Evidence from the Garner task. *Cognition and Emotion*, 18(2), 209-231.

Glauert, R., Rhodes, G., Byrne, S., Fink, B., & Grammer, K. (2009). Body dissatisfaction and the effects of perceptual exposure on body norms and ideals. *International Journal of Eating Disorders*, 42(5), 443-452.

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *Journal of cognitive neuroscience*, 19(11), 1803-1814.

Golan, T., Bentin, S., DeGutis, J. M., Robertson, L. C., & Harel, A. (2014). Association and dissociation between detection and discrimination of objects of expertise: Evidence from visual search. *Attention, Perception, & Psychophysics*, 76(2), 391-406.

Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature neuroscience*, 4(8), 845.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1), 20-25.

Goshen-Gottstein, Y., & Ganel, T. (2000). Repetition priming for familiar and unfamiliar faces in a sex-judgment task: Evidence for a common route for the processing of sex and identity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1198.

Graham, R., & LaBar, K. S. (2007). Garner interference reveals dependencies between emotional expression and gaze in face perception. *Emotion*, 7(2), 296.

Greene, M. R. (2013). Statistics of high-level scene context. *Front. Psychol.* 4, 777.

Greven, I. M., & Ramsey, R. (2017). Person perception involves functional integration between the extrastriate body area and temporal pole. *Neuropsychologia*, 96, 52-60.

Greven, I. M., & Ramsey, R. (2017)a. Person perception involves functional integration between the extrastriate body area and temporal pole. *Neuropsychologia*, 96, 52-60.

Greven, I. M., & Ramsey, R. (2017)b. Neural network integration during the perception of in-group and out-group members. *Neuropsychologia*, 106, 225-235.

Greven, I. M., Downing, P. E., & Ramsey, R. (2016). Linking person perception and person knowledge in the human brain. *Social cognitive and affective neuroscience*, 11(4), 641-651.

Greven, I. M., Downing, P. E., & Ramsey, R. (2018). Neural networks supporting social evaluation of bodies based on body shape. *Social neuroscience*, 14(3), 328-344.

Grezes, J., Pichon, S., & De Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage*, 35(2), 959-967.

Grosbras, M. H., & Paus, T. (2005). Brain networks involved in viewing angry hands or faces. *Cerebral Cortex*, 16(8), 1087-1096.

Grosbras, M. H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. *Human brain mapping*, 33(2), 431-454.

Gross, C. G. (1992). Representation of visual stimuli in inferior temporal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 335(1273), 3-10.

Gross, C. G., Rocha-Miranda, C. D., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of neurophysiology*, 35(1), 96-111.

Grossman, E. D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision research*, 41(10-11), 1475-1482.

Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron*, 35(6), 1167-1175.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of cognitive neuroscience*, 12(5), 711-720.

- Gu, X., Liu, X., Guise, K. G., Naidich, T. P., Hof, P. R., & Fan, J. (2010). Functional dissociation of the frontoinsula and anterior cingulate cortices in empathy for pain. *Journal of Neuroscience*, 30(10), 3739-3744.
- Hadjikhani, N., & de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Current Biology*, 13(24), 2201-2205.
- Harris, A., & Aguirre, G. K. (2008). The representation of parts and wholes in face-selective cortex. *Journal of Cognitive Neuroscience*, 20(5), 863-878.
- Harris, A., & Aguirre, G. K. (2010). Neural tuning for face wholes and parts in human fusiform gyrus revealed by fMRI adaptation. *Journal of Neurophysiology*, 104(1), 336-345.
- Harris, A., Vyas, D. B., & Reed, C. L. (2016). Holistic processing for bodies and body parts: New evidence from stereoscopic depth manipulations. *Psychonomic bulletin & review*, 23(5), 1513-1519.
- Harris, D. A., & Ciaramitaro, V. M. (2016). Interdependent mechanisms for processing gender and emotion: the special status of angry male faces. *Frontiers in psychology*, 7, 1046.
- Harry, B. B., Umla-Runge, K., Lawrence, A. D., Graham, K. S., & Downing, P. E. (2016). Evidence for integrated visual face and body representations in the anterior temporal lobes. *Journal of cognitive neuroscience*, 28(8), 1178-1193.
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of personality and social psychology*, 78(1), 81.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and social psychology Review*, 10(1), 47-66.
- Hasselmo, M. E., Rolls, E. T., Baylis, G. C., & Nalwa, V. (1989). Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental Brain Research*, 75(2), 417-429.
- Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150366.
- Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, 37(1), 25.
- Haugh, S. S., Hoffman, C. D., & Cowan, G. (1980). The eye of the very young beholder: Sex typing of infants by young children. *Child development*, 598-600.

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

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

Hehman, E., Carpinella, C. M., Johnson, K. L., Leitner, J. B., & Freeman, J. B. (2014). Early processing of gendered facial cues predicts the electoral success of female politicians. *Social Psychological and Personality Science*, 5(7), 815-824.

Hehman, E., Stoller, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Processes & Intergroup Relations*, 18(3), 384-401.

Henriksson, L., Mur, M. and Kriegeskorte, N. (2019). Rapid invariant encoding of scene layout in human OPA. *Neuron*. 04, 014.

Heron-Delaney, M., Wirth, S., & Pascalis, O. (2011). Infants' knowledge of their own species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1571), 1753-1763.

Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop out effect for faces. *Vision research*, 45(13), 1707-1724.

Hershler, O., & Hochstein, S. (2009). The importance of being expert: Top-down attentional control in visual search with photographs. *Attention, Perception, & Psychophysics*, 71(7), 1478-1486.

Hess, U., Adams, R. B., Grammer, K., & Kleck, R. E. (2009). Face gender and emotion expression: Are angry women more like men?. *Journal of Vision*, 9(12), 19-19.

Hickey, C., & Peelen, M. V. (2015). Neural mechanisms of incentive salience in naturalistic human vision. *Neuron*, 85(3), 512-518.

Hickey, C., & van Zoest, W. (2012). Reward creates oculomotor salience. *Current Biology*, 22(7), R219-R220.

Hickey, C., Kaiser, D., & Peelen, M. V. (2015). Reward guides attention to object categories in real-world scenes. *Journal of Experimental Psychology: General*, 144(2), 264.

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual review of psychology*, 38(1), 369-425.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of experimental social psychology*, 13(2), 141-154.

Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: The role of shape and colour. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 261(1362), 367-373.

Hill, M. Q., Streuber, S., Hahn, C. A., Black, M. J., & O'Toole, A. J. (2016). Creating body shapes from verbal descriptions by linking similarity spaces. *Psychological science*, 27(11), 1486-1497.

Hock, A., Kangas, A., Zieber, N., & Bhatt, R. S. (2015). The development of sex category representation in infancy: Matching of faces and bodies. *Developmental psychology*, 51(3), 346.

Hodzic, A., Kaas, A., Muckli, L., Stirn, A., & Singer, W. (2009)a. Distinct cortical networks for the detection and identification of human body. *Neuroimage*, 45(4), 1264-1271.

Hodzic, A., Muckli, L., Singer, W., & Stirn, A. (2009)b. Cortical responses to self and others. *Human brain mapping*, 30(3), 951-962.

Hoff, H., & Pötzl, O. (1937). Über eine optisch-agnostische Störung des „Physiognomie-Gedächtnisses“. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, 159(1), 367-395.

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

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, 23(1), 65-74.

Hortensius, R., Terburg, D., Morgan, B., Stein, D. J., van Honk, J., & de Gelder, B. (2017). The Basolateral Amygdalae and Frontotemporal network functions for threat perception. *Eneuro*, 4(1).

Hsu, S. M., & Young, A. (2004). Adaptation effects in facial expression recognition. *Visual Cognition*, 11(7), 871-899.

Hu, Y., Parde, C. J., Hill, M. Q., Mahmood, N., & O'Toole, A. J. (2018). First Impressions of Personality Traits From Body Shapes. *Psychological science*, 29(12), 1969-1983.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148(3), 574-591.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1), 106-154.

Hugenberg, K. (2005). Social categorization and the perception of facial affect: target race moderates the response latency advantage for happy faces. *Emotion*, 5(3), 267.

Hugenberg, K., & Sczesny, S. (2006). On wonderful women and seeing smiles: Social categorization moderates the happy face response latency advantage. *Social Cognition*, 24(5), 516-539.

Hulleman, J., Te Winkel, W., & Boselie, F. (2000). Concavities as basic features in visual search: Evidence from search asymmetries. *Perception & Psychophysics*, 62(1), 162-174.

Hummel, D., Grabhorn, R., & Mohr, H. M. (2012). Body-shape adaptation cannot be explained by adaptation to narrow and wide rectangles. *Perception*, 41(11), 1315-1322.

Humphrey, N. K. (1976). The social function of intellect. In *Growing points in ethology* (pp. 303-317). Cambridge University Press.

Humphreys, G. W., Donnelly, N., & 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(2), 173-181.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences*, 96(16), 9379-9384.

Jamovi project. (2019). jamovi (Version 0.9). <https://www.jamovi.org/>.

Jastorff, J., & Orban, G. A. (2009). Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing. *Journal of Neuroscience*, 29(22), 7315-7329.

Jellema, T., & Perrett, D. I. (2003). Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion?. *Neuropsychologia*, 41(13), 1728-1737.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics*, 14(2), 201-211.

Johnson, K. L., & Tassinari, L. G. (2005). Perceiving sex directly and indirectly: Meaning in motion and morphology. *Psychological Science*, 16(11), 890-897.

Johnson, K. L., Iida, M., & Tassinari, L. G. (2012). Person (mis) perception: Functionally biased sex categorization of bodies. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749), 4982-4989.

Johnson, K. L., Lurye, L. E., & Tassinari, L. G. (2010). Sex categorization among preschool children: Increasing utilization of sexually dimorphic cues. *Child development*, 81(5), 1346-1355.

Johnson, K., & Shiffrar, M. (Eds.). (2012). *People watching: Social, perceptual, and neurophysiological studies of body perception*. Oxford University Press.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Harvard University Press.

Johnston, W. A., & Hawley, K. J. (1994). Perceptual inhibition of expected inputs: The key that opens closed minds. *Psychonomic Bulletin & Review*, 1(1), 56-72.

Johnstone, L. T., & Downing, P. E. (2017). Dissecting the visual perception of body shape with the Garner selective attention paradigm. *Visual Cognition*, 25(4-6), 507-523.

Julian, J. B., Ryan, J., Hamilton, R. H. and Epstein, R. A. (2016). The occipital place area is causally involved in representing environmental boundaries during navigation. *Curr. Biol.* 26, 1104-1109.

Kable, J. W., & Chatterjee, A. (2006). Specificity of action representations in the lateral occipitotemporal cortex. *Journal of Cognitive Neuroscience*, 18(9), 1498-1517.

Kaiser, D., Azzalini, D. C., & Peelen, M. V. (2016). Shape-independent object category responses revealed by MEG and fMRI decoding. *Journal of neurophysiology*, 115(4), 2246-2250.

Kaiser, D., Quek, G. L., Cichy, R. M., & Peelen, M. V. (2019). Object vision in a structured world. *Trends in cognitive sciences*.

Kaiser, D., Stein, T., & Peelen, M. V. (2014). Object grouping based on real-world regularities facilitates perception by reducing competitive interactions in visual cortex. *Proceedings of the National Academy of Sciences*, 111(30), 11217-11222.

Kaiser, D., Strnad, L., Seidl, K. N., Kastner, S., & Peelen, M. V. (2013). 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(1), 82-90.

Kamps, F. S., Julian, J. B., Kubilius, J., Kanwisher, N. and Dilks, D. D. (2016). The occipital place area represents the local elements of scenes. *Neuroimage* 132, 417-424.

Kana, R. K., & Travers, B. G. (2012). Neural substrates of interpreting actions and emotions from body postures. *Social cognitive and affective neuroscience*, 7(4), 446-456.

Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di psicologia*, 49(1), 7-30.

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

Karnadewi, Fika, and Ottmar V. Lipp. "The processing of invariant and variant face cues in the Garner Paradigm." *Emotion* 11, no. 3 (2011): 563.

Kaul, C., Ratner, K. G., & Van Bavel, J. J. (2012). Dynamic representations of race: processing goals shape race decoding in the fusiform gyri. *Social cognitive and affective neuroscience*, 9(3), 326-332.

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. In *Advances in experimental social psychology* (Vol. 55, pp. 1-80). Academic Press.

Kelley, H. H. (1950). The warm-cold variable in first impressions of persons. *Journal of personality*, 18(4), 431-439.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. Guilford Press.

Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: a social relations analysis. *Psychological bulletin*, 102(3), 390.

Kiani, R., Esteky, H., Mirpour, K., & Tanaka, K. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *Journal of neurophysiology*, 97(6), 4296-4309.

Kitada, R., Yoshihara, K., Sasaki, A. T., Hashiguchi, M., Kochiyama, T., & Sadato, N. (2014). The brain network underlying the recognition of hand gestures in the blind: the supramodal role of the extrastriate body area. *Journal of Neuroscience*, 34(30), 10096-10108.

Kite, M. E., Deaux, K., & Haines, E. L. (2008). Gender stereotypes. *Psychology of women: A handbook of issues and theories*, 2, 205-236.

Knoblich, G. E., Thornton, I. M., Grosjean, M. E., & Shiffrar, M. E. (2006). *Human body perception from the inside out: Advances in visual cognition*. Oxford University Press.

Kok, P., Brouwer, G. J., van Gerven, M. A., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *Journal of Neuroscience*, 33(41), 16275-16284.

Kok, P., Failing, M. F. and de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. *J. Cogn. Neurosci.* 26, 1546-1554.

Kok, P., Jehee, J. F. and de Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron* 75, 265-270.

Kok, P., Jehee, J. F., & De Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, 75(2), 265-270.

- Kok, P., Mostert, P. and de Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proc. Natl. Acad. Sci. U. S. A.* 114, 10473-10478.
- Kovács, G., Zimmer, M., Bankó, É., Harza, I., Antal, A., & Vidnyánszky, Z. (2005). Electrophysiological correlates of visual adaptation to faces and body parts in humans. *Cerebral Cortex*, 16(5), 742-753.
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & psychophysics*, 21(6), 575-580.
- Kramer, R. S., Young, A. W., Day, M. G., & Burton, A. M. (2017). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review*, 124(2), 115.
- Kret, M. E., & De Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, 50(7), 1211-1221.
- Kret, M. E., Pichon, S., Grèzes, J., & de Gelder, B. (2011). Similarities and differences in perceiving threat from dynamic faces and bodies. An fMRI study. *Neuroimage*, 54(2), 1755-1762.
- Kristjánsson, Á., & Tse, P. U. (2001). Curvature discontinuities are cues for rapid shape analysis. *Perception & Psychophysics*, 63(3), 390-403.
- Kumar, S., Popivanov, I. D., & Vogels, R. (2017). Transformation of visual representations across ventral stream body-selective patches. *Cerebral Cortex*, 29(1), 215-229.
- Lange, J., & Lappe, M. (2006). A model of biological motion perception from configural form cues. *Journal of Neuroscience*, 26(11), 2894-2906.
- Langton, S. R., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, 107(1), 330-342.
- Larson, C. L., Aronoff, J., & Stearns, J. J. (2007). The shape of threat: Simple geometric forms evoke rapid and sustained capture of attention. *Emotion*, 7(3), 526.
- Latinus, M., & Belin, P. (2011). Human voice perception. *Current Biology*, 21(4), R143-R145.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human perception and performance*, 21(3), 451.
- Lawrence, M. A. (2016). ez: Easy Analysis and Visualization of Factorial Experiments. R package version 4.4-0. <https://rdr.io/cran/ez/>.

Lawson, R. P., Clifford, C. W., & Calder, A. J. (2009). About turn: The visual representation of human body orientation revealed by adaptation. *Psychological Science*, 20(3), 363-371.

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

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

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General*, 129(4), 559.

Li, X., Cave, K. R., & Wolfe, J. M. (2008). Kanizsa-type subjective contours do not guide attentional deployment in visual search but line termination contours do. *Perception & Psychophysics*, 70(3), 477-488.

Light, L. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, 5(3), 212.

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). The many faces of research on face perception.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature neuroscience*, 5(9), 910.

Lupyan, G. and Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Curr. Direct. Psychol. Sci.* 24, 279-284.

Lupyan, G. and Thompson-Schill, S. L. (2012). The evocative power of words: activation of concepts by verbal and nonverbal means. *J. Exp. Psychol. Gen.* 141, 170-186.

Macdonald, S. N., & Culham, J. C. (2015). Do human brain areas involved in visuomotor actions show a preference for real tools over visually similar non-tools?. *Neuropsychologia*, 77, 35-41.

Mack, A., & Rock, I. (1998). *Inattention blindness*. MIT press.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin*, 107(3), 401.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual review of psychology*, 51(1), 93-120.

Macrae, C. N., & Martin, D. (2007). A boy primed Sue: Feature-based processing and person construal. *European Journal of Social Psychology*, 37(5), 793-805.

Macrae, C. N., & Quadflieg, S. (2010). Perceiving people.

Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M., & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology*, 76(2), 200.

Macrae, C. N., Stangor, C., & Milne, A. B. (1994). Activating social stereotypes: A functional analysis. *Journal of Experimental Social Psychology*, 30(4), 370-389.

Makowski, (2018). *The psycho Package: An Efficient and Publishing-Oriented Workflow for Psychological Science*. Journal of Open Source Software, 3(22), 470. <https://doi.org/10.21105/joss.00470>

Malcolm, G. L., Groen, I. I. A. and Baker, C. I. (2016). Making sense of real-world scenes. *Trends. Cogn. Sci.* 20, 843-856.

Markus, H., & Zajonc, R. B. (1985). The cognitive perspective in social psychology. *Handbook of social psychology*, 1, 137-230.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 207(1167), 187-217.

Martin, A., & Weisberg, J. (2003). Neural foundations for understanding social and mechanical concepts. *Cognitive neuropsychology*, 20(3-6), 575-587.

Martin, D., & Macrae, C. N. (2007). A face with a cue: Exploring the inevitability of person categorization. *European Journal of Social Psychology*, 37(5), 806-816.

Martin, D., & Macrae, C. N. (2010). Processing style and person recognition: Exploring the face inversion effect. *Visual Cognition*, 18(2), 161-170.

Martinez, L., Falvello, V. B., Aviezer, H., & Todorov, A. (2016). Contributions of facial expressions and body language to the rapid perception of dynamic emotions. *Cognition and Emotion*, 30(5), 939-952.

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, 14(9), 744-751.

Matsuno, T., & Fujita, K. (2018). Body inversion effect in monkeys. *PloS one*, 13(10), e0204353.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in cognitive sciences*, 6(6), 255-260.

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological review*, 90(3), 215.

McBain, R., Norton, D., & Chen, Y. (2009). Females excel at basic face perception. *Acta psychologica*, 130(2), 168-173.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in cognitive sciences*, 7(7), 293-299.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 102(3), 419.

McGugin, R. W., Gatenby, J. C., Gore, J. C., & Gauthier, I. (2012). High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance. *Proceedings of the National Academy of Sciences*, 109(42), 17063-17068.

McGugin, R. W., McKeeff, T. J., Tong, F., & Gauthier, I. (2011). Irrelevant objects of expertise compete with faces during visual search. *Attention, Perception, & Psychophysics*, 73(2), 309-317.

McMahon, D. B., & Olson, C. R. (2007). Repetition suppression in monkey inferotemporal cortex: relation to behavioral priming. *Journal of neurophysiology*, 97(5), 3532-3543.

Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, 102(45), 16518-16523.

Meteyard, L. and Holmes, N. P. (2018). TMS SMART - Scalp mapping of annoyance ratings and twitches caused by Transcranial Magnetic Stimulation. *J. Neurosci. Methods* 299, 34-44.

Minnebusch, D. A., & Daum, I. (2009). Neuropsychological mechanisms of visual face and body perception. *Neuroscience & Biobehavioral Reviews*, 33(7), 1133-1144.

Minnebusch, D. A., Keune, P. M., Suchan, B., & Daum, I. (2010). Gradual inversion affects the processing of human body shapes. *Neuroimage*, 49(3), 2746-2755.

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L.,

Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). Face perception during early infancy. *Psychological Science*, 10(5), 419-422.

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). Face perception during early infancy. *Psychological Science*, 10(5), 419-422.

Montagne, B., Kessels, R. P., Frigerio, E., de Haan, E. H., & Perrett, D. I. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity?. *Cognitive processing*, 6(2), 136-141.

Moro, V., Urgesi, C., Pernigo, S., Lanteri, P., Pazzaglia, M., & Aglioti, S. M. (2008). The neural basis of body form and body action agnosia. *Neuron*, 60(2), 235-246.

Morrison, I., Peelen, M. V., & Downing, P. E. (2006). The sight of others' pain modulates motor processing in human cingulate cortex. *Cerebral cortex*, 17(9), 2214-2222.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychological review*, 98(2), 164.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of cognitive neuroscience*, 9(5), 555-604.

Murray, J. E., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science*, 11(6), 492-496.

Myers, A., & Sowden, P. T. (2008). Your hand or mine? The extrastriate body area. *Neuroimage*, 42(4), 1669-1677.

Nakayama, K. (1994). James J. Gibson: An appreciation.

Neisser, U. (1963). Decision-time without reaction-time: Experiments in visual scanning. *The American journal of psychology*, 76(3), 376-385.

Nestor, A., & Tarr, M. J. (2008). Gender recognition of human faces using color. *Psychological Science*, 19(12), 1242-1246.

Ng, M., Ciaramitaro, V. M., Anstis, S., Boynton, G. M., & Fine, I. (2006). Selectivity for the configural cues that identify the gender, ethnicity, and identity of faces in human cortex. *Proceedings of the National Academy of Sciences*, 103(51), 19552-19557.

Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment.

Noyes, E., Hill, M. Q., & O'Toole, A. J. (2018). Face recognition ability does not predict person identification performance: using individual data in the interpretation of group results. *Cognitive research: principles and implications*, 3(1), 23.

- O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, 401(6753), 584.
- Oliva, A. and Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International J. Computer Vision* 42, 145-175.
- Olson, J. M., Roese, N. J., & Zanna, M. P. (1996). Expectancies.
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087-11092.
- Pallier, C. (2002). Computing discriminability and bias with the R software. URL <http://www.pallier.org/ressources/aprime/aprime> (Last viewed March 18, 2016).
- Palumbo, R., D'Ascenzo, S., & Tommasi, L. (2015). Cross-category adaptation: exposure to faces produces gender aftereffects in body perception. *Psychological research*, 79(3), 380-388.
- Palumbo, R., Laeng, B., & Tommasi, L. (2013). Gender-specific aftereffects following adaptation to silhouettes of human bodies. *Visual Cognition*, 21(1), 1-12.
- Panichello, M. F., Cheung, O. S. and Bar, M. (2012). Predictive feedback and conscious visual experience. *Front. Psychol.* 3, 620.
- Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion effect. *Psychological science*, 28(3), 369-379.
- Parry, F. M., Young, A. W., Shona, J., Saul, M., & Moss, A. (1991). Dissociable face processing impairments after brain injury. *Journal of Clinical and Experimental Neuropsychology*, 13(4), 545-558.
- Peelen, M. V. and Kastner, S. (2014). Attention in the real world: toward understanding its neural basis. *Trends Cogn. Sci.* 18, 242-250.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of neurophysiology*, 93(1), 603-608.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, 8(8), 636.
- Peelen, M. V., & Downing, P. E. (2017). Category selectivity in human visual cortex: Beyond visual object recognition. *Neuropsychologia*, 105, 177-183.
- Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences*, 108(29), 12125-12130.

- Peelen, M. V., & Kastner, S. (2014). Attention in the real world: toward understanding its neural basis. *Trends in cognitive sciences*, 18(5), 242-250.
- Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *Journal of Neuroscience*, 30(30), 10127-10134.
- Peelen, M. V., Atkinson, A. P., Andersson, F., & Vuilleumier, P. (2007). Emotional modulation of body-selective visual areas. *Social Cognitive and Affective Neuroscience*, 2(4), 274-283.
- Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, 460(7251), 94.
- 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(6), 815-822.
- Peelen, Marius V., Chenxi He, Zaizhu Han, Alfonso Caramazza, and Yanchao Bi. "Nonvisual and visual object shape representations in occipitotemporal cortex: evidence from congenitally blind and sighted adults." *Journal of Neuroscience* 34, no. 1 (2014): 163-170.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical transactions of the royal society of London. Series B: Biological sciences*, 335(1273), 23-30.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental brain research*, 47(3), 329-342.
- Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. *Human neurobiology*, 3(4), 197-208.
- Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. A. (2005). Specificity of regions processing biological motion. *European Journal of Neuroscience*, 21(10), 2864-2875.
- Picci, G., & Scherf, K. S. (2016). From caregivers to peers: puberty shapes human face perception. *Psychological Science*, 27(11), 1461-1473.
- Pichon, S., de Gelder, B., & Grezes, J. (2008). Emotional modulation of visual and motor areas by dynamic body expressions of anger. *Social neuroscience*, 3(3-4), 199-212.
- Pick, A. (1922). Störung der Orientierung am eigenen Körper. *Psychological Research*, 1(1), 303-318.

- Pinsk, M. A., DeSimone, K., Moore, T., Gross, C. G., & Kastner, S. (2005). Representations of faces and body parts in macaque temporal cortex: a functional MRI study. *Proceedings of the National Academy of Sciences*, 102(19), 6996-7001.
- Pitcher, D., Charles, L., Devlin, J. T., Walsh, V. and Duchaine, B. (2009). Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Curr. Biol.* 19, 319-324.
- 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(4), 2356-2363.
- Pitcher, D., Goldhaber, T., Duchaine, B., Walsh, V. and Kanwisher, N. (2012). Two critical and functionally distinct stages of face and body perception. *J. Neurosci.* 32, 15877-15885.
- Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The gender stereotyping of emotions. *Psychology of Women Quarterly*, 24(1), 81-92.
- Pomerantz, J. R., & Garner, W. R. (1973). Stimulus configuration in selective attention tasks. *Perception & psychophysics*, 14(3), 565-569.
- Ponzo, M. (1911). *Recherches sur la localisation des sensations tactiles et des sensations douloureuses*. Administration des Archives italiennes de biologie.
- Popivanov, I. D., Jastorff, J., Vanduffel, W., & Vogels, R. (2012). Stimulus representations in body-selective regions of the macaque cortex assessed with event-related fMRI. *Neuroimage*, 63(2), 723-741.
- Popivanov, I. D., Jastorff, J., Vanduffel, W., & Vogels, R. (2014). Heterogeneous single-unit selectivity in an fMRI-defined body-selective patch. *Journal of Neuroscience*, 34(1), 95-111.
- Popivanov, I. D., Schyns, P. G., & Vogels, R. (2016). Stimulus features coded by single neurons of a macaque body category selective patch. *Proceedings of the National Academy of Sciences*, 113(17), E2450-E2459.
- Pourtois, G., Peelen, M. V., Spinelli, L., Seeck, M., & Vuilleumier, P. (2007). Direct intracranial recording of body-selective responses in human extrastriate visual cortex. *Neuropsychologia*, 45(11), 2621-2625.
- Proklova, D., Kaiser, D., & Peelen, M. V. (2016). Disentangling representations of object shape and object category in human visual cortex: The animate–inanimate distinction. *Journal of cognitive neuroscience*, 28(5), 680-692.
- Puri, A. M., Wojciulik, E. and Ranganath, C. (2009). Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Res.* 1301, 89-99.

Quadflieg, S., & Koldewyn, K. (2017). The neuroscience of people watching: how the human brain makes sense of other people's encounters. *Annals of the New York Academy of Sciences*, 1396(1), 166-182.

Quadflieg, S., & Macrae, C. N. (2011). Stereotypes and stereotyping: What's the brain got to do with it?. *European Review of Social Psychology*, 22(1), 215-273.

Quadflieg, S., Flannigan, N., Waiter, G. D., Rossion, B., Wig, G. S., Turk, D. J., & Macrae, C. N. (2011). Stereotype-based modulation of person perception. *Neuroimage*, 57(2), 549-557.

Quinn, K. A., & Macrae, C. N. (2011). The face and person perception: Insights from social cognition. *British Journal of Psychology*, 102(4), 849-867.

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, 31(9), 1109-1121.

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

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

Ralph, M. A. L., Jefferies, E., Patterson, K. and Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nat. Rev. Neurosci.* 18, 42-55.

Ramsey, J. L., Langlois, J. H., & Marti, N. C. (2005). Infant categorization of faces: Ladies first. *Developmental Review*, 25(2), 212-246.

Ramsey, R. (2018). Neural integration in body perception. *Journal of cognitive neuroscience*, 30(10), 1442-1451.

Ramsey, R., & Hamilton, A. F. D. C. (2010). Understanding actors and object-goals in the human brain. *Neuroimage*, 50(3), 1142-1147.

Ramsey-Rennels, J. L., & Langlois, J. H. (2006). Infants' differential processing of female and male faces. *Current Directions in Psychological Science*, 15(2), 59-62.

Reed, C. L., Bukach, C. M., Garber, M., & McIntosh, D. N. (2018). It's Not All About the Face: Variability Reveals Asymmetric Obligatory Processing of Faces and Bodies in Whole-Body Contexts. *Perception*, 47(6), 626-646.

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological science*, 14(4), 302-308.

Reed, C. L., Stone, V. E., Grubb, J. D., & McGoldrick, J. E. (2006). Turning configural processing upside down: part and whole body postures. *Journal of Experimental Psychology: Human Perception and Performance*, 32(1), 73.

Reeder, R. R., & Peelen, M. V. (2013). The contents of the search template for category-level search in natural scenes. *Journal of Vision*, 13(3), 13-13.

Reeder, R. R., Stein, T., & Peelen, M. V. (2016). Perceptual expertise improves category detection in natural scenes. *Psychonomic bulletin & review*, 23(1), 172-179.

Reeder, R. R., van Zoest, W., & Peelen, M. V. (2015). Involuntary attentional capture by task-irrelevant objects that match the search template for category detection in natural scenes. *Attention, Perception, & Psychophysics*, 77(4), 1070-1080.

Rennels, J. L., Juvrud, J., Kayl, A. J., Asperholm, M., Gredebäck, G., & Herlitz, A. (2017). Caregiving experience and its relation to perceptual narrowing of face gender. *Developmental psychology*, 53(8), 1437.

Rensink, R. A., & Cavanagh, P. (2004). The influence of cast shadows on visual search. *Perception*, 33(11), 1339-1358.

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision research*, 46(18), 2977-2987.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces?. *Cognition*, 47(1), 25-57.

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive psychology*, 19(4), 473-497.

Rhodes, G., Jeffery, L., Boeing, A., & Calder, A. J. (2013). Visual coding of human bodies: perceptual aftereffects reveal norm-based, opponent coding of body identity. *Journal of Experimental Psychology: Human Perception and Performance*, 39(2), 313.

Rhodes, G., Robbins, R., Jaquet, E., McKone, E., Jeffery, L., & Clifford, C. W. (2005). Adaptation and face perception: How aftereffects implicate norm-based coding of faces. *Fitting the mind to the world: Adaptation and after-effects in high-level vision*, 213-240.

Rice, A., Phillips, P. J., & O'Toole, A. (2013)b. The role of the face and body in unfamiliar person identification. *Applied Cognitive Psychology*, 27(6), 761-768.

Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. J. (2013)a. Unaware person recognition from the body when face identification fails. *Psychological Science*, 24(11), 2235-2243.

Ro, T., Friggel, A., & Lavie, N. (2007). Attentional biases for faces and body parts. *Visual Cognition*, 15(3), 322-348.

Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological science*, 12(1), 94-99.

Robbins, R. A., & Coltheart, M. (2012). The effects of inversion and familiarity on face versus body cues to person recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), 1098.

Rolls, E. T. (1984). Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Human neurobiology*, 3(4), 209-222.

Rolls, E. T., Baylis, G. C., & Leonard, C. M. (1985). Role of low and high spatial frequencies in the face-selective responses of neurons in the cortex in the superior temporal sulcus in the monkey. *Vision research*, 25(8), 1021-1035.

Rosch, E. H. (1972). Probabilities, sampling, and ethnographic method: The case of Dani colour names. *Man*, 7(3), 448-466.

Rosch, E. H. (1973). Natural categories. *Cognitive psychology*, 4(3), 328-350.

Rosenholtz, R. (2001). Search asymmetries? What search asymmetries?. *Perception & Psychophysics*, 63(3), 476-489.

Rosenthal, R., & DePaulo, B. M. (1979). Sex differences in eavesdropping on nonverbal cues. *Journal of Personality and social Psychology*, 37(2), 273.

Rosenthal, R., Archer, D., Hall, J. A., DiMatteo, M. R., & Rogers, P. L. (1979). Measuring sensitivity to nonverbal communication: The PONS test. In *Nonverbal Behavior* (pp. 67-98). Academic Press.

Rossi, S., Hallett, M., Rossini, P. M. and Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin. Neurophys.* 120, 2008-2039.

Rossi, S., Hallett, M., Rossini, P. M. and Pascual-Leone, A. (2011). Screening questionnaire before TMS: an update. *Clin. Neurophys.* 122, 1686.

Rossini, P. M. et al. (2015). Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: basic principles and procedures for routine clinical and research application. An updated report from an IFCN Committee. *Clin. Neurophys.* 126, 1071-1107.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces?. *Behavioral and cognitive neuroscience reviews*, 1(1), 63-75.

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(11), 2381-2395.

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature neuroscience*, 8(1), 107.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological review*, 89(1), 60.

Sacchett, C., & Humphreys, G. W. (1992). Calling a squirrel a squirrel but a canoe a wigwam: A category-specific deficit for artefactual objects and body parts. *Cognitive Neuropsychology*, 9(1), 73-86.

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

Said, C. P., Dotsch, R., & Todorov, A. (2010). The amygdala and FFA track both social and non-social face dimensions. *Neuropsychologia*, 48(12), 3596-3605.

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: a defense of functional localizers. *Neuroimage*, 30(4), 1088-1096.

Saxe, R., Jamal, N., & Powell, L. (2005). My body or yours? The effect of visual perspective on cortical body representations. *Cerebral Cortex*, 16(2), 178-182.

Schupp, H. T., & Renner, B. (2011). The implicit nature of the anti-fat bias. *Frontiers in Human Neuroscience*, 5, 23.

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience*, 25(47), 11055-11059.

Schweinberger, S. R., & Burton, A. M. (2011). Person perception 25 years after Bruce and Young (1986): An introduction.

Schweinberger, S. R., & Soukup, G. R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology: Human perception and performance*, 24(6), 1748.

Schweinberger, S. R., Burton, A. M., & Kelly, S. W. (1999). Asymmetric dependencies in perceiving identity and emotion: Experiments with morphed faces. *Perception & Psychophysics*, 61(6), 1102-1115.

Schweinberger, S. R., Klos, T., & Sommer, W. (1995). Covert face recognition in prosopagnosia: A dissociable function?. *Cortex*, 31(3), 517-529.

Schweinberger, Stefan R., Christoph Casper, Nadine Hauthal, Jürgen M. Kaufmann, Hideki Kawahara, Nadine Kloth, David MC Robertson, Adrian P. Simpson, and Romi Zäske. "Auditory adaptation in voice perception." *Current Biology* 18, no. 9 (2008): 684-688.

- Schwoebel, J., & Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of cognitive neuroscience*, 17(4), 543-553.
- Sekunova, A., Black, M., Parkinson, L., & Barton, J. J. (2013). Viewpoint and pose in body-form adaptation. *Perception*, 42(2), 176-186.
- Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2008). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences*, 276(1656), 575-584.
- Sergent, J., & Signoret, J. L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral cortex*, 2(5), 375-388.
- Sergent, J., Ohta, S., & MACDONALD, B. (1992). Functional neuroanatomy of face and object processing: a positron emission tomography study. *Brain*, 115(1), 15-36.
- Silvanto, J., Muggleton, N. and Walsh, V. (2008). State-dependency in brain stimulation studies of perception and cognition. *Trends Cogn. Sci.* 12, 447-454.
- Simanova, I., Francken, J. C., de Lange, F. P. and Bekkering, H. (2016). Linguistic priors shape categorical perception. *Lang. Cogn. Neurosci.* 31, 159-165.
- Simion, F., Leo, I., Turati, C., Valenza, E., & Dalla Barba, B. (2007). How face specialization emerges in the first months of life. *Progress in brain research*, 164, 169-185.
- Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current opinion in neurobiology*, 13(2), 144-149.
- Singh, D. (1993). Adaptive significance of female physical attractiveness: role of waist-to-hip ratio. *Journal of personality and social psychology*, 65(2), 293.
- Skinner, B. F. (1953). Some contributions of an experimental analysis of behavior to psychology as a whole. *American Psychologist*, 8(2), 69.
- Slaughter, V., Heron, M., & Sim, S. (2002). Development of preferences for the human body shape in infancy. *Cognition*, 85(3), B71-B81.
- Slaughter, V., Heron, M., Jenkins, L., Tilse, E., Müller, U., & Liebermann, D. (2004). Origins and early development of human body knowledge. *Monographs of the Society for Research in Child Development*, i-113.
- Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences*, 107(46), 20099-20103.
- Sorger, B., Goebel, R., Schiltz, C., & Rossion, B. (2007). Understanding the functional neuroanatomy of acquired prosopagnosia. *Neuroimage*, 35(2), 836-852.

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, 62(2), 207.

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, 62(2), 207.

Steeves, J. K., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., ... & 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(4), 594-609.

Stein, T., Sterzer, P., & Peelen, M. V. (2012). Privileged detection of conspecifics: Evidence from inversion effects during continuous flash suppression. *Cognition*, 125(1), 64-79.

Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport*, 15(5), 777-780.

Stephen, I. D., Bickersteth, C., Mond, J., Stevenson, R. J., & Brooks, K. R. (2016). No effect of featural attention on body size aftereffects. *Frontiers in psychology*, 7, 1223.

Stokes, M., Thompson, R., Nobre, A. C. and Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19569-19574.

Stolier, R. M., & Freeman, J. B. (2016). Neural pattern similarity reveals the inherent intersection of social categories. *Nature neuroscience*, 19(6), 795.

Stolier, R. M., & Freeman, J. B. (2017). A neural mechanism of social categorization. *Journal of Neuroscience*, 37(23), 5711-5721.

Sturman, D., Stephen, I. D., Mond, J., Stevenson, R. J., & Brooks, K. R. (2017). Independent Aftereffects of Fat and Muscle: Implications for neural encoding, body space representation, and body image disturbance. *Scientific reports*, 7, 40392.

Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, 13(3), 283-286.

Summerfield, C. and de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* 15, 745-756.

Summerfield, C. and Egnér, T. (2009). Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* 13, 403-409.

Summerfield, C. et al. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science* 314, 1311-1314.

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, 15(11), 745-756.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Eger, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature neuroscience*, 11(9), 1004.

Sun, J. Y., & Perona, P. (1996)a. Preattentive perception of elementary three-dimensional shapes. *Vision Research*, 36(16), 2515-2529.

Sun, J., & Perona, P. (1996)b. Early computation of shape and reflectance in the visual system. *Nature*, 379(6561), 165.

Suzuki, S., & Cavanagh, P. (1995). Facial organization blocks access to low-level features: An object inferiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, 21(4), 901.

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

Tamietto, M., Geminiani, G., Genero, R., & de Gelder, B. (2007). Seeing fearful body language overcomes attentional deficits in patients with neglect. *Journal of cognitive neuroscience*, 19(3), 445-454.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly journal of experimental psychology*, 46(2), 225-245.

Taylor, J. C., & Downing, P. E. (2011). Division of labor between lateral and ventral extrastriate representations of faces, bodies, and objects. *Journal of Cognitive Neuroscience*, 23(12), 4122-4137.

Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *Journal of neurophysiology*, 98(3), 1626-1633.

Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2010). fMRI-adaptation studies of viewpoint tuning in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, 103(3), 1467-1477.

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature neuroscience*, 10(4), 505.

- Thierry, G., Pegna, A. J., Dodds, C., Roberts, M., Basan, S., & Downing, P. (2006). An event-related potential component sensitive to images of the human body. *Neuroimage*, 32(2), 871-879.
- Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception*.
- Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual review of psychology*, 66, 519-545.
- Tong, F., & Nakayama, K. (1999). Robust representations for faces: evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4), 1016.
- Tong, F., & Nakayama, K. (1999). Robust representations for faces: evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4), 1016.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological review*, 95(1), 15.
- Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, 114(3), 285.
- Troiani, V., Stigliani, A., Smith, M. E. and Epstein, R. A. (2014). Multiple object properties drive scene-selective regions. *Cereb. Cortex* 24, 883-897.
- Troje, N. F., Sadr, J., Geyer, H., & Nakayama, K. (2006). Adaptation aftereffects in the perception of gender from biological motion. *Journal of vision*, 6(8), 7-7.
- Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological review*, 93(3), 239.
- Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. (2003). Faces and objects in macaque cerebral cortex. *Nature neuroscience*, 6(9), 989.
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, 311(5761), 670-674.
- Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences*, 105(49), 19514-19519.
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1), 71-86.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological review*, 80(1), 1.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157), 1124-1131.

Uher, R., Murphy, T., Friederich, H. C., Dalgleish, T., Brammer, M. J., Giampietro, V., ... & Campbell, I. C. (2005). Functional neuroanatomy of body shape perception in healthy and eating-disordered women. *Biological psychiatry*, 58(12), 990-997.

Urgesi, C., Berlucchi, G. and Aglioti, S. M. (2004). Magnetic stimulation of extrastriate body area impairs visual processing of nonfacial body parts. *Curr. Biol.* 14, 2130-2134.

Urgesi, C., Calvo-Merino, B., Haggard, P., & Aglioti, S. M. (2007). Transcranial magnetic stimulation reveals two cortical pathways for visual body processing. *Journal of Neuroscience*, 27(30), 8023-8030.

Urgesi, C., Candidi, M., Ionta, S., & Aglioti, S. M. (2007). Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. *Nature neuroscience*, 10(1), 30.

Valentin, D., Abdi, H., O'Toole, A. J., & Cottrell, G. W. (1994). Connectionist models of face processing: A survey. *Pattern recognition*, 27(9), 1209-1230.

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

Valentine, T., & Bruce, V. (1986). The effects of distinctiveness in recognising and classifying faces. *Perception*, 15(5), 525-535.

Valyear, K. F., & Culham, J. C. (2010). Observing learned object-specific functional grasps preferentially activates the ventral stream. *Journal of Cognitive Neuroscience*, 22(5), 970-984.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: a functional magnetic resonance imaging investigation. *Psychological science*, 19(11), 1131-1139.

van de Riet, W. A., Grezes, J., & de Gelder, B. (2009). Specific and common brain regions involved in the perception of faces and bodies and the representation of their emotional expressions. *Social neuroscience*, 4(2), 101-120.

van den Hurk, J., Van Baelen, M., & de Beeck, H. P. O. (2017). Development of visual category selectivity in ventral visual cortex does not require visual experience. *Proceedings of the National Academy of Sciences*, 114(22), E4501-E4510.

Van Essen, D. C., & Maunsell, J. H. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in neurosciences*, 6, 370-375.

van Koningsbruggen, M. G., Peelen, M. V. and Downing, P. E. (2013). A causal role for the extrastriate body area in detecting people in real-world scenes. *J. Neurosci.* 33, 7003-7010.

VandenBos, G. R. (2007). *APA dictionary of psychology*. American Psychological Association.

Vangeneugden, J., Peelen, M. V., Tadin, D., & Battelli, L. (2014). Distinct neural mechanisms for body form and body motion discriminations. *Journal of Neuroscience*, 34(2), 574-585.

Vocks, S., Busch, M., Grönemeyer, D., Schulte, D., Herpertz, S., & Suchan, B. (2010). Neural correlates of viewing photographs of one's own body and another woman's body in anorexia and bulimia nervosa: an fMRI study. *Journal of psychiatry & neuroscience: JPN*, 35(3), 163.

Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik* (Vol. 9). Voss.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in cognitive sciences*, 9(12), 585-594.

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia*, 45(1), 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(3), 829-841.

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. *Nature neuroscience*, 7(11), 1271.

Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cerebral Cortex*, 4(5), 509-522.

Wagemans, J., De Winter, J., de Beeck, H. O., Ploeger, A., Beckers, T., & Vanroose, P. (2008). Identification of everyday objects on the basis of silhouette and outline versions. *Perception*, 37(2), 207-244.

Walbrin, J., & Koldewyn, K. (2019). Dyadic interaction processing in the posterior temporal cortex. *NeuroImage*, 198, 296-302.

Wallbott, H. G. (1998). Bodily expression of emotion. *European journal of social psychology*, 28(6), 879-896.

Wang, D., Kristjansson, A., & Nakayama, K. (2005). Efficient visual search without top-down or bottom-up guidance. *Perception & Psychophysics*, 67(2), 239-253.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, 107(3), 829-853.

Watson, T. L., Otsuka, Y., & Clifford, C. W. G. (2016). Who are you expecting? Biases in face perception reveal prior expectations for sex and age. *Journal of Vision*, 16(3), 5-5.

Webster, M. A., & MacLeod, D. I. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1571), 1702-1725.

Webster, M. A., & Maclin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic bulletin & review*, 6(4), 647-653.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, 428(6982), 557.

Wenger, M. J., & Townsend, J. T. (2001). Faces as gestalt stimuli: Process characteristics. *Computational, geometric, and process perspectives on facial cognition*, 229-284.

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis* (New York: Springer).

Wickham, H., François, R., L. H. and, K. M. (2019). *dplyr: A Grammar of Data Manipulation*. R package version 0.8.0.1. <https://dplyr.tidyverse.org/>.

Wiggett, A. J., & Downing, P. E. (2011). Representation of action in occipito-temporal cortex. *Journal of Cognitive Neuroscience*, 23(7), 1765-1780.

Willems, S., Vrancken, L., Germeys, F., & Verfaillie, K. (2014). Holistic processing of human body postures: evidence from the composite effect. *Frontiers in psychology*, 5, 618.

Wilson, J. P., Hugenberg, K., & Rule, N. O. (2017). Racial bias in judgments of physical size and formidability: From size to threat. *Journal of personality and social psychology*, 113(1), 59.

Winkler, C., & Rhodes, G. (2005). Perceptual adaptation affects attractiveness of female bodies. *British Journal of Psychology*, 96(2), 141-154.

Winston, J. S., Gottfried, J. A., Kilner, J. M., & Dolan, R. J. (2005). Integrated neural representations of odor intensity and affective valence in human amygdala. *Journal of Neuroscience*, 25(39), 8903-8907.

Winston, J. S., Henson, R. N. A., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of neurophysiology*, 92(3), 1830-1839.

Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I., & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, 45(1), 195-206.

Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & psychophysics*, 63(3), 381-389.

Xu, X., & Biederman, I. (2010). Loci of the release from fMRI adaptation for changes in facial expression, identity, and viewpoint. *Journal of Vision*, 10(14), 36-36.

Yamaguchi, M. K., Hirukawa, T., & Kanazawa, S. (2013). Judgment of gender through facial parts. *Perception*, 42(11), 1253-1265.

Yang, H., Susilo, T., & Duchaine, B. (2014). The anterior temporal face area contains invariant representations of face identity that can persist despite the loss of right FFA and OFA. *Cerebral Cortex*, 26(3), 1096-1107.

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

Yon, D., Gilbert, S. J., de Lange, F. P. and Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nat. Commun.* 9, 4288.

Young, A. W., Hellawell, D., & Hay, D. C. (2013). Configurational information in face perception. *Perception*, 42(11), 1166-1178.

Young, A. W., & Bruce, V. (2011). Understanding person perception. *British journal of psychology*, 102(4), 959-974.

Young, A. W., Newcombe, F., Haan, E. H. D., Small, M., & Hay, D. C. (1993). Face perception after brain injury: Selective impairments affecting identity and expression. *Brain*, 116(4), 941-959.

Yovel, G., & Belin, P. (2013). A unified coding strategy for processing faces and voices. *Trends in cognitive sciences*, 17(6), 263-271.

Yovel, G., & O'Toole, A. J. (2016). Recognizing people in motion. *Trends in cognitive sciences*, 20(5), 383-395.

Yovel, G., Pelc, T., & Lubetzky, I. (2010). It's all in your head: Why is the body inversion effect abolished for headless bodies?. *Journal of Experimental Psychology: Human Perception and Performance*, 36(3), 759.

Zeki, S. M. (1978). Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *The Journal of Physiology*, 277(1), 273-290.

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision research*, 41(23), 2979-2994.

Zieber, N., Bhatt, R. S., Hayden, A., Kangas, A., Collins, R., & Bada, H. (2010). Body representation in the first year of life. *Infancy*, 15(5), 534-544.

Zieber, N., Kangas, A., Hock, A., & Bhatt, R. S. (2014). Infants' perception of emotion from body movements. *Child Development*, 85(2), 675-684.

Zimmermann, M., Verhagen, L., de Lange, F. P. and Toni, I. (2016). The extrastriate body area computes desired goal states during action planning. *eNeuro* 3, e0020-16.2016 1-13.

Supplementary Materials



fear



angry

Towards a model of human body perception



sad



happy

Supplementary Figure S3.1. Shows all the pictures of body emotional expressions in experiment 1-3.

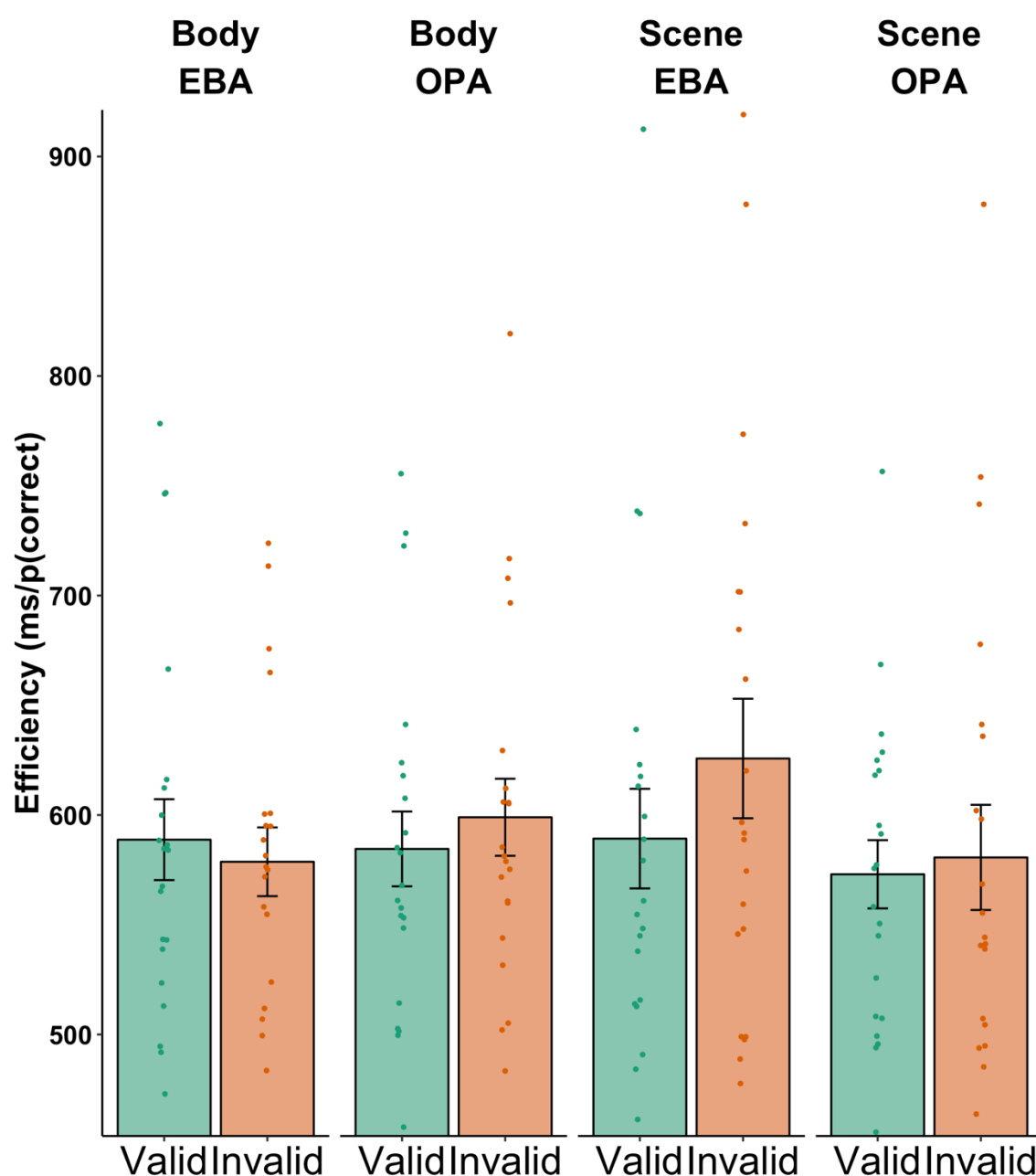


Figure S4.1. Impact of TMS over extrastriate category-selective regions on cue-driven stimulus expectations, related to Figure 2. Mean efficiency scores across participants (RT / p(correct)) are plotted in relation to cue validity, task, and stimulation site. Bars indicate mean values; error bars SE of the mean (including within- and between-participants variance); individual points reflect scores for each participant.

Towards a model of human body perception

ID	Task	Area	x	y	z
1-1	Body	EBA	40.00	-73.00	-2.00
1-2	Body	EBA	48.00	-65.00	0.00
1-3	Body	EBA	51.00	-70.00	9.00
1-4	Body	EBA	46.00	-74.00	11.00
1-5	Body	EBA	50.00	-65.00	6.00
1-6	Body	EBA	52.00	-65.00	-1.00
1-7	Body	EBA	52.00	-59.00	-1.00
1-8	Body	EBA	48.00	-72.00	-2.00
1-9	Body	EBA	45.00	-59.00	3.00
1-10	Body	EBA	50.00	-69.00	-1.00
1-11	Body	EBA	44.00	-74.00	6.00
1-12	Body	EBA	44.00	-71.00	10.00
1-13	Body	EBA	49.00	-73.00	3.00
1-14	Body	EBA	40.00	-70.00	0.00
1-15	Body	EBA	51.00	-78.00	3.00
1-16	Body	EBA	53.00	-70.00	-2.00
1-17	Body	EBA	49.00	-76.00	-5.00
1-18	Body	EBA	50.00	-67.00	1.00
1-19	Body	EBA	43.00	-80.00	2.00
1-20	Body	EBA	44.00	-70.00	2.00
1-21	Body	EBA	52.00	-65.00	1.00

Towards a model of human body perception

2-1	Scene	EBA	50.00	-68.00	1.00
2-2	Scene	EBA	44.00	-64.00	5.00
2-3	Scene	EBA	45.00	-80.00	-2.00
2-4	Scene	EBA	45.00	-76.00	-1.00
2-5	Scene	EBA	52.00	-70.00	-1.00
2-6	Scene	EBA	41.00	-80.00	16.00
2-7	Scene	EBA	47.00	-54.00	2.00
2-8	Scene	EBA	43.00	-81.00	-1.00
2-9	Scene	EBA	43.00	-76.00	6.00
2-10	Scene	EBA	47.00	-73.00	-7.00
2-11	Scene	EBA	46.00	-76.00	-2.00
2-12	Scene	EBA	56.00	-58.00	-8.00
2-13	Scene	EBA	50.00	-69.00	-3.00
2-14	Scene	EBA	57.00	-73.00	2.00
2-15	Scene	EBA	49.00	-75.00	3.00
2-16	Scene	EBA	55.00	-71.00	1.00
2-17	Scene	EBA	55.00	-79.00	5.00
2-18	Scene	EBA	46.00	-78.00	7.00
2-19	Scene	EBA	49.00	-70.00	2.00
2-20	Scene	EBA	51.00	-77.00	3.00
2-21	Scene	EBA	49.00	-70.00	1.00
1-1	Body	OPA	36.00	-83.00	15.00

Towards a model of human body perception

1-2	Body	OPA	26.00	-75.00	23.00
1-3	Body	OPA	31.00	-71.00	16.00
1-4	Body	OPA	33.00	-75.00	34.00
1-5	Body	OPA	35.00	-76.00	8.00
1-6	Body	OPA	38.00	-73.00	25.00
1-7	Body	OPA	30.00	-78.00	7.00
1-8	Body	OPA	29.00	-77.00	21.00
1-9	Body	OPA	37.00	-75.00	23.00
1-10	Body	OPA	29.00	-78.00	25.00
1-11	Body	OPA	23.00	-83.00	28.00
1-12	Body	OPA	32.00	-77.00	15.00
1-13	Body	OPA	40.00	-76.00	17.00
1-14	Body	OPA	25.00	-81.00	15.00
1-15	Body	OPA	29.00	-76.00	20.00
1-16	Body	OPA	36.00	-77.00	20.00
1-17	Body	OPA	35.00	-80.00	23.00
1-18	Body	OPA	38.00	-73.00	24.00
1-19	Body	OPA	40.00	-83.00	21.00
1-20	Body	OPA	37.00	-76.00	18.00
1-21	Body	OPA	35.00	-80.00	20.00
2-1	Scene	OPA	32.00	-75.00	23.00
2-2	Scene	OPA	30.00	-81.00	15.00

Towards a model of human body perception

2-3	Scene	OPA	36.00	-84.00	15.00
2-4	Scene	OPA	26.00	-79.00	23.00
2-5	Scene	OPA	31.00	-78.00	9.00
2-6	Scene	OPA	24.00	-82.00	30.00
2-7	Scene	OPA	45.00	-71.00	25.00
2-8	Scene	OPA	35.00	-86.00	24.00
2-9	Scene	OPA	29.00	-81.00	18.00
2-10	Scene	OPA	35.00	-72.00	5.00
2-11	Scene	OPA	33.00	-78.00	19.00
2-12	Scene	OPA	35.00	-79.00	17.00
2-13	Scene	OPA	37.00	-79.00	15.00
2-14	Scene	OPA	46.00	-85.00	15.00
2-15	Scene	OPA	37.00	-87.00	18.00
2-16	Scene	OPA	42.00	-82.00	25.00
2-17	Scene	OPA	34.00	-90.00	26.00
2-18	Scene	OPA	36.00	-86.00	24.00
2-19	Scene	OPA	34.00	-83.00	20.00
2-20	Scene	OPA	38.00	-86.00	15.00
2-21	Scene	OPA	30.00	-81.00	21.00

Table S4.1. MNI coordinates of peak locations identified in independent functional localizers and targeted for TMS, related to Figure 1. Separate groups of participants performed each task.

Accuracy (% Correct)				
Condition	<i>M</i>	<i>M</i>		<i>SD</i>
		95% CI	[LL, UL]	
Body-EBA-Valid	0.89	[0.86, 0.92]		0.07
Body-EBA-Invalid	0.91	[0.88, 0.94]		0.07
Body-OPA-Valid	0.91	[0.89, 0.94]		0.06
Body-OPA-Invalid	0.92	[0.89, 0.94]		0.06
Scene-EBA-Valid	0.90	[0.87, 0.93]		0.06
Scene-EBA-Invalid	0.87	[0.83, 0.91]		0.09
Scene-OPA-Valid	0.90	[0.87, 0.93]		0.07
Scene-OPA-Invalid	0.91	[0.87, 0.94]		0.08
RTs in ms				
Condition	<i>M</i>	<i>M</i>		<i>SD</i>
		95% CI	[LL, UL]	
Body-EBA-Valid	523.64	[483.07, 564.21]		89.13
Body-EBA-Invalid	528.91	[489.37, 568.46]		86.88
Body-OPA-Valid	534.07	[497.21, 570.94]		80.98
Body-OPA-Invalid	549.09	[509.24, 588.95]		87.56
Scene-EBA-Valid	525.21	[491.73, 558.70]		73.56
Scene-EBA-Invalid	537.98	[501.82, 574.14]		79.44
Scene-OPA-Valid	514.87	[483.97, 545.76]		67.87
Scene-OPA-Invalid	522.44	[482.62, 562.26]		87.48
Efficiency (RT in ms / % Correct)				

Condition	<i>M</i>		
	<i>M</i>	95% CI	<i>SD</i>
		[LL, UL]	
Body-EBA-Valid	588.76	[550.33, 627.19]	84.43
Body-EBA-Invalid	578.68	[546.03, 611.34]	71.74
Body-OPA-Valid	584.54	[548.99, 620.10]	78.10
Body-OPA-Invalid	598.97	[562.36, 635.57]	80.42
Scene-EBA-Valid	589.24	[541.93, 636.56]	103.95
Scene-EBA-Invalid	625.75	[568.94, 682.55]	124.79
Scene-OPA-Valid	573.00	[540.57, 605.43]	71.24
Scene-OPA-Invalid	580.68	[530.71, 630.64]	109.76

Table S4.2. Means and standard deviations for accuracy, response times, and efficiency in each condition, related to Figure 2. Naming convention: Task-Stimulation Site-Condition; e.g. Body-EBA-Valid describes performance on valid trials from the body task under stimulation over extrastriate body area (EBA). *M* and *SD* represent mean and standard deviation, respectively. LL and UL indicate the lower and upper limits of the 95% confidence interval for the mean, respectively.