The perceptual integrality of sex and age: understanding the functional organisation of face processing

Aitken, Paul

Award date:
2024

Awarding institution:
Bangor University

PRIFYSGOL
BANGOR
UNIVERSITY

The perceptual integrality of sex and age: understanding the functional organisation of face processing

Paul George Aitken

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, 2023

The perceptual integrality of sex and age

‘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.

I confirm that I am submitting this work with the agreement of my Supervisor(s).’

‘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.

Rwy'n cadarnhau fy mod yn cyflwyno'r gwaith hwn gyda chytundeb fy Ngoruchwyliwr (Goruchwylwyr)

Preface and acknowledgements

I have been accused of many things in my life, but never of being a man of few words. Failing to keep this thesis brief, I wanted at least to keep this part brief. In the end, I decided not to break the habit of a lifetime. What would make this section too long – even by my standards – would be to single out by name every person that has directly or indirectly contributed to the current work. To have too many people to thank personally is a gift for which I am most grateful.

Before thanking people for their part in it, a brief word about the work. It would have been completely different if it were not for the 2020 coronavirus pandemic. The work that constitutes the current thesis began towards the end of the summer of 2021, with everything carried out until that point being shelved. Despite the stress of changing direction around the half-way mark, I am very grateful for it.

Between my masters and PhD, I had come to suspect that perception must involve some bidirectional process of integration and separation – of putting some sensory and semantic elements together and keeping others apart. I did not know about the long and rich theoretical and experimental history of the problem. It was only under the pressure of adaptation to material circumstances that I became aware of, and ultimately initiated into, the school of thought which deals with perceptual integrality. I feel deeply humbled to number myself amongst those concerned with it, and hope to have been able to contribute to its study in a meaningful way. I would have liked more time to worry about it - but I would likely have wasted it.

I want to thank my family first – some of whom, unfortunately, are no longer here to celebrate in this achievement, but they've always felt near nonetheless. It is no exaggeration to say that I wouldn't be in this position without my mother, father, and brother – who have provided me with more than I can say. They have been unquestioningly supportive and respectful of me even when they haven't understood or agreed with me. I love them very much and am incredibly grateful to them.

I am thankful to my friends at home and abroad for keeping in touch with me, coming to see me, inviting me to see them and for keeping me abreast of what is happening in the world beyond

Bangor. I feel blessed to have so many true friends in my life and am excited to be a bigger part of what they do next.

I would like to thank Professor Paul Downing for his attentive and adaptable supervision. His approach to his work balances careful, analytical and systematic practice with curiosity, humility and passion – his was precisely the type of academic and scientific training that I needed. When I started working with Paul, I dwelt almost entirely in the clouds – but with his help I've been able to tentatively lower one foot to the ground. Without his supervision I'm doubtful that I would ever have finished this thesis, I am grateful for his patience with me.

I'm also thankful to the rest of the faculty at the psychology department at Bangor University for their warm, welcoming acceptance, and for treating me like a junior colleague rather than a subordinate. Such an attitude towards myself and the other students has offered a candid and exemplary insight into the reality of academic work. I'm glad to have had the opportunity to be here, and wish everyone the very best for the future. I'd particularly like to thank Kami Koldewyn for employing me to assist on her project at a time when I really needed it,

I want to thank my fellow PGR students, past and present, for sharing in all of my successes and defeats. I have always felt enabled and encouraged by these colleagues to bring my ideas to fruition – even the bad ones. More importantly, I have had countless of the types of conversations that remind me why I do this kind of work. I see no impediment to the continuation and development of good, intentional, considerate and wonderful science and philosophy with such people amongst my cohort. I am inspired by them, and carry their dedication with me.

I think that's everyone… No, wait. Hector. I am more grateful to you than you know. Getting to know you over the past few years has been like coming to terms with a part of myself. I never question whether what we've been through has been worth it. Without you to share in this experience with me, it would have cost me more than double and given me less than half. I can't wait to see what happens next.

And, of course, thanks go above all to God. There's one for my materialist-reductionist colleagues.

This work is dedicated to
Áine Mullan

Summary

We spend our lives looking at other people’s faces. Although we seem to perceive them instantaneously as wholes, the face is made up of multiple different sources of overlapping information – about different socially-relevant characteristics of the face such as sex, age, race, emotional expression and eye gaze direction. The ease with which we perceive all of this information belies the immense computational complexity involved in processing it.

Each dimension relates to a different source of information about the person. Questions remain as to how information from these sources interact or interfere with one another in processing. The field of social vision has emerged to tackle such problems, integrating research in vision science and in social psychology around the idea that perceiving information about other people relies on the complex, dynamic interplay of different types of information. This information is either sensory or semantic; relating to that which is encoded through the senses, or to meaningful information about a given object that we have already encountered and represented in memory respectively.

In this thesis, I aim to extend the social vision approach to the problem of perceptual integration in face processing by exploring the under-investigated perceptual relationship between sex and age. I utilise five well-established experimental paradigms from the cognitive psychology literature to test participants' capacity to selectively attend to sex and age cues across several sets of faces. Namely, I carried out a Garner speeded classification task, an Eriksen flanker task, a face priming study, a word priming study, and a task-switching study.

Taken together, the studies suggest that sex and age are asymmetrically integrated in human face processing at sensory and semantic levels of representation, such that sex information interferes with age judgements more so than vice versa. I propose that a speed-of-processing account (SoPA) of dimension discriminability (Melara & Mounts, 1993) - by which it is assumed that the less discriminable (harder to process) dimension interferes more with the more discriminable (easier to process) dimension than vice versa - cannot fully account for the current results. I suggest an alternative to this account which incorporates a perceptual mechanism to account for differences in dimensional discriminability and asymmetric integration between sex

and age: that each involves a different “depth-of-processing”, such that age perception can take place using relatively superficial or “shallow” featural processing, whilst sex perception typically requires relatively complex or “deep” configural processing. I call this the depth-of-processing account (DoPA)

Contents

Chapter 1. The Problem of Vision

1.1 Conceptual and Historical Issues in Visual Perception

> *I have yet to see any problem, however complicated, which when you looked at it the right way did not become still more complicated*
>
> *Poul Anderson - 1957*

Francis Crick described the embarrassed pauses that might arise at the dinner table when he would answer “seeing”, to a guest when asked what he was currently working on.

> *My questioner is wondering why there should be any difficulty about something as simple as seeing. After all, we open our eyes and there the world is, large and clear, full of objects in vivid Technicolor, without our having to make any appreciable effort. It all seems so delightfully easy, so what can be the problem?*
>
> *Crick - 1990*

The central mistake in the naive intuition about vision is that it is easy; what is closer to the truth is that human beings are extremely good at it (Marr, 1982). Marr pointed out that scientists didn’t realise how complex of a problem vision was until they started trying to program computers to do it. Not long before trying to embody the visual process in an artificial medium, mainstream experimental psychology tended to view it as essentially simple, passive and objective rather than complex, active and subjective - a process which is “so incredibly complex that it seems almost a miracle that we can do it at all” (Palmer, 1999).

Whilst technological advances and information and computational theory helped to advance our view of vision, there were numerous other factors and intermediary steps in the transition of vision (and perception, more generally) away from simplicity and towards complexity. Crucially, there has been acknowledgment of and research into internal states and mental contents such as goals, memory, motivation, emotion, thought, and attention, and the recognition that such states, contents and processes affect perceptual experience.

What follows is a brief historical overview of how mainstream experimental psychology has come to embrace and better understand the complexity of vision and the visual system, and of cognitive and perceptual functioning more generally; as well as the transition of focus in mainstream psychology away from behaviour and towards cognition in the mid-twentieth century - permitting the emergence of the interdisciplinary field of vision science, one of the many cognitive sciences. This overview is far from exhaustive, but provides historical context for the contemporary view of vision and the visual system which informs and directs the current work.

1.1.2 Empirical perception and the "New Look"

The assumption that vision is "simple" may be rooted in psychological ideas about perception from the previous century. Before 1940, perceptual research was typically restricted to psychophysical investigations of the constancy hypothesis, the idea that percepts would mirror stimuli in a one-to-one manner - that the same stimulus would always produce the same sensation and the same changes to a stimulus would always produce the same corresponding changes in consciousness. Perception in experimental psychology was seen as essentially a "pure" and passive phenomenon; with the individual regarded as an essentially interchangeable "complex recording instrument", the unobservable internal states of which having no bearing on the outcome of the perceptual process (Wertz, 1983). In the first half of the twentieth century, experimental psychology was wholly empirical, and has been characterised as abiding by the "robot mode of human behaviour" (Von Bertalanffy, 1967)

Although Helmholtz (1867, 2009) had advocated the idea that there was a gap between optical stimulation and conscious perception which needed to be filled by hidden "assumptions" in order to reach perceptual "conclusions"; and although philosophers such as Kant (1781, 1908) had argued against the idea of knowledge on the grounds of pure empiricism - emphasising the necessity of a priori knowledge in perception - mainstream experimental psychology was not yet able to cope with the idea that perception could vary as a function of the individual. "Experimental psychology looked upon this problem of projective perception, because of its complexity and experimental unmanageability, as unfit for scientific analysis" (Werner & Wapner, 1949). Some researchers recognised that, despite being easier to study, mainstream empirical and behaviourist approaches were not true to the complexity of conscious awareness (Bruner & Postman, 1949; Werner & Wapner, 1949).

The stubborn empiricist outlook began to change with the emergence of the "New Look" movement, described after the fact by Jerome Bruner - one of its key architects - as "a new mentalism on its way to becoming the Cognitive Revolution" (Bruner, 1992). This school of thought insisted that perception could not be understood independently of observers, that two different observers may "see" two things differently, and that the same stimulus might have different perceptual effects when presented repeatedly to the same observer. The basic axiom of New Look research was that perception results from a prepared (eigenstellt) organism rather than a passive or neutral one (Bruner & Postman, 1949), that we are involved in the construction of our perception.

> *Perception was not, in the positivist sense, a mere neutral registration of what was "out there,", but was, rather, an activity affected by other concurrent processes of thought, memory, and so on.*
>
> *Bruner - 1992*

As is the case with many theoretical advances, progress occurred in tandem with technological advances which permitted new lines of research to be carried out. The emergence of the tachistoscope - an instrument used to expose the eye to visual objects very briefly - permitted researchers to discover, for instance, that the recognition threshold of words presented with equal frequencies differed depending on the semantic value of those words to the observer (Postman, Bruner & McGinnies, 1948). this shows that perception was affected by pre-recognition semantic qualities of a stimulus, not just its physical attributes.

In another such experiment (Bruner & Postman, 1949), participants were shown ordinary playing cards and "trick" cards (impossible cards such as a red four of clubs or a black four of hearts). Participants were asked to identify the cards which were initially presented briefly, but their duration increased as the experiment progressed. Response times for trick cards were not significantly different than for ordinary cards at short presentation durations, and the participant would either get the answer right without noticing the incongruity, or would get the answer wrong in a manner which corresponded with the incongruity (identifying a red four of hearts as a four of spades, for example). Increasing display durations caused an increase in response times, as participants became more aware of the incongruity, until eventually the duration was long enough for them to identify the issue. Such an experiment highlights the power of expectation in perceptual decision making, with participants responding in terms of their previous experiences and expectations rather than the visual stimulation itself.

Early tachistoscopic experiments such as these showed - within the accepted parameters of psychological experimentation - that internal states affected perceptual mechanics, that what we already know and understand can affect what we see and the decisions that we make based on what we see. According to Bruner (1992), however, the progression of the New Look movement was adversely affected by a schism in focus between those interested in cognition for its own sake, and those using the new experimental and theoretical advances to investigate and reify psychodynamic theory.

Such psychodynamically oriented research helped critics of the New Look to level scrutiny against the entire endeavour at a time when Behaviourism was still the dominant school of thought, and when psychodynamic research was considered inherently unscientific (Wertz, 1983). They tended to either argue that the studies measured something other than perception, or that they lacked proper methodological control. There were also failures of replication which challenged the empirical findings. The schism in research direction in the New Look continued until 1956 - a year of “crucial importance” (Eysenck & Keane, 2015), and “the mythical birthday of the cognitive revolution” (Bruner, 1992).

1.1.3 “The Cognitive Revolution”

The cognitive revolution “brought the mind back into experimental psychology”, after mainstream experimentalists had “proposed to redefine psychology as the science of behaviour” (Miller, 2003). With this, a step was taken towards making experimental psychology more realistic, and closer to a genuine study of subjective reality. As well as having the necessary “mentalism” in place thanks to the now faltering New Look school in perceptual psychology and in line with a confluence of external factors, the cognitive revolution is widely thought to have begun at the symposium on information theory held at MIT in the summer of 1956 (Eysenck & Keane, 2015;

Miller, 2003; Solso, 1991). George Miller, who presented his now classic research which suggested that the information processing capacity of working memory constrains retrieval to around the “Magic Number Seven”, later wrote of the symposium at MIT:

> *I went away from the Symposium with a strong conviction, more intuitive than rational, that human experimental psychology, theoretical linguistics and computer simulation of rational cognitive processes were all pieces of a larger whole, and that the future would see progressive elaboration and coordination of their shared concerns.*
>
> *Miller - 1979*

Some of the factors which contributed to the beginning of the cognitive revolution included the ostensible inadequacy of behaviourism to fully account for the diversity of human behaviour (Solso, 1991). American psychologists such as Bruner and Miller were emboldened to forge their own path by international scholars whose work was less ensconced in the terminology of behaviourism. In the UK, Bartlett (1949) recognised social influences on memory, such that participants tended to translate the culturally unfamiliar aspects of folk stories into more culturally familiar terms in their recall.

In Switzerland, Piaget’s developmental work (e.g. Piaget & Inhelder, 1948) suggested amongst other things that children could simulate the internal states of others. In Russia, Luria (1932) wrote of the monism of mind and brain - rejecting the distinction between the two, seeing psychological processes as inalienable from neuroanatomy. Work such as this, and the keen followers that such researchers had amassed, helped the Americans to develop a new psychology, one which could “free itself from behaviourism, thus restoring cognition to scientific respectability” (Miller, 2003).

Other factors which influenced the onset of the cognitive revolution included the advent of computer sciences in the wake of World War II. There was the theory of servo motors (engines which required internal feedback), information theory and, signal-detection theory (a computational framework for describing the extraction of signal from noise (Green & Swets, 1966)) - which had been used to detect and decode enemy radio signals during the war. From 1940, computers capable of working with complex numbers were developed (Rees, 1980), and the realisation that unseen, internal mechanisms and processes were necessary for the completion of such tasks began to permeate the broader scientific consciousness. Of particular import were cybernetics (Wiener, 1950) and artificial intelligence (McCarthy, Minsky, Rochester & Shannon, 1955), which raised questions about the structure and function of memory, including its encoding, storage and retrieval.

There was a growing feeling amongst several disciplines including linguistics (Chomsky, 1956) that collaboration was needed in order to address certain major, trans-disciplinary problems. Researchers were quick to adopt and develop what appeared to be a shared language through which to engage in this collaboration - the languages of computer science and information theory (e.g. Shannon, 1948). It was the MIT symposium in September of 1956 which made it clear that these languages were experimentally and theoretically applicable to problems of the mind. The operations of the mind came to be analogised and measured in terms of information-processing computations. By 1960, the application of information-processing based ideas about the mind was being called cognitive studies, information-processing psychology, or cognitive science.

With this, the migration from a largely behaviourist to a largely cognitivist view in mainstream psychology was underway, transforming what was meant by “the unconscious”. No longer regarded as a hazily defined, near-mystical concept of relevance only in cases of psychopathology, it is now generally accepted that cognitive processes operating outwith the scope of perceptual awareness bear upon the outcomes of that awareness (see Kihlstrom, 1987 for a review). Today, to suggest otherwise - that all cognitive and perceptual processes must be available to consciousness - seems “downright barmy” (Bruner, 1992).

Theories about internal processes, the language of information theory, and the technology which permitted experiments into those processes and couched in that language led over time to the development four major branches of cognitive science - all of which are concerned with hypothesised internal processes such as attention, learning, memory and reasoning. According to Eysenck & Keane (2015), these are cognitive psychology, which can be described in terms of the study of behaviour through the lens of internal processes using experimentation and observation; cognitive neuropsychology, the study of cases of impaired functioning, brain disease and brain damage to elucidate and define information processing theories of the mind; cognitive neuroscience, which involves some combination of behavioural and brain-based data - acquired through technologies such as functional magnetic resonance imaging (fMRI), trans-cranial magnetic stimulation (TMS) and electroencephalogram (EEG); and computational cognitive science, by which computational models are developed in order to simulate hypothesised cognitive operations.

Although the precise definitions of these approaches are open to debate, they tell us about the behaviour of people with healthy brains and damaged brains, at the level of behaviour and the brain, and permit us to simulate their operations in order to test their plausibility. Through the adoption of the language of information theory and its evolution and adaptation within the cognitive sciences these four major approaches have elucidated a great deal about the nature of perceptual and cognitive systems and functioning. Alongside *information theory*, two other major theoretical constructs have emerged which characterise current views of vision and of the visual system - *representation* and *hierarchical organisation*.

The following section explores the history and conceptual bases of information, representation and hierarchical organisation in more detail, with more specific behavioural, computational and neuroanatomical evidence for the operation of these concepts being outlined throughout the thesis. The explication and use of these concepts situates the current work firmly within the tradition of cognitive psychology. A broad understanding of these concepts, their assumptions and how they are proposed to relate to one another frames an understanding of the empirical work detailed below.

1.2 The Three Pillars of Contemporary Vision Science

1.2.1 Information

> *It is much easier to talk about information than it is to say what you are talking about*
>
> *Dretske - 1983*

The language of information processing and computer science seems so ubiquitous in modern society that most people have an intuitive understanding of their concepts - of information, signals, encoding and decoding, processing, computation, storage and transmission, inputs and outputs. Although these terms do have technical definitions, Garner (1962) points out that they are not particularly at odds with the lay definitions.

Information theory and its concepts are largely credited to work by Shannon (1948), who proved that increasing the amount of information in a signal does not necessitate an increase in noise within the channel through which the signal is communicated. Although the concepts with which he dealt were already present in the burgeoning communication sciences, Shannon's genius was, according to Gleick (2011), in simplifying, mathematising and "purifying" those concepts - making information into something quantifiable and measurable - in the same way as Newton did with the previously vague, "unscientific" and immeasurable definitions of "force", "motion" and "mass". Shannon and Newton took vague terms and made them more precise, facilitating their use in the statement of scientific observations (Chalmers, 1999).

History will determine whether Shannon's quantification of these concepts has had the same impact on humanity as that of Newton, but it has certainly revolutionised the world of science and technology. "If you want to understand life…" wrote Dawkins (1996) "think about information technology". According to Loewenstein (1999), information "Connotes a cosmic principle of organisation and order, and it provides an exact measurement of that", for Wheeler (1999), information is the basis of "Every it - every particle, every field of force, even the spacetime continuum itself". These statements suppose that information is simultaneously the underpinning for life and the universe as well as the means of understanding them, but what is it?

Defining information seems surprisingly difficult:

> *An exploration of "information" runs into immediate difficulties since information has to do with becoming informed, with the reduction of ignorance and of uncertainty, it is ironic that the term "information" is itself ambiguous and used in different ways*
>
> *Buckland - 1991*

For Shannon (1948), information is a measurement of the average amount of data arriving at a receiver from a source, it is that which is transmitted across a channel from the source to the receiver. In terms of visual perception, the individual might be the receiver, a visual object the source, and the channel across which the information is transmitted is the visual system. There are many ways to define the term more technically, although it is best used to define it with respect to the specific science and theory that the definition is adopted to explain (Capurro & Hjørland, 2005). With that in mind, I will employ the definition of information supplied by Wendell Garner, who's work forms a large part of the scientific and theoretical backbone of the current work.

> *Information is something we get when some person or machine tells us something we didn't know before. We may read a book, have a question answered or simply observe or listen to the world. By these various acts we acquire knowledge or information about the world - at least in most cases.*
>
> *Garner - 1962*

For Garner, information is the inverse of uncertainty - it is that which reduces uncertainty.

> *Thus information exists in a message or communication only if there is an a priori uncertainty about what the message will be… The amount of information is determined by the amount of the uncertainty - or, more exactly, it is determined by the amount by which the uncertainty has been reduced.*
>
> *Ibid.*

Miller (1956) equated "information" to "the variance we have been talking about for years". In Miller's view, the benefit of speaking in terms of information rather than the specific type of variance observed in a given experiment (pertaining to pounds, volts, responses etc.) is that the former is a "dimensionless quality", a quantity unto its own that needn't be transformed between cases in order to make sense. If there is a lot of variance, there is a lot of information when we make an observation because we cannot readily predict what is going to happen - the more variance, the less we can predict (the higher the uncertainty) and thus the more information we attain when the observation is made. On the other hand, when we have very little variance, the outcome of a single observation contains little information - we can predict the outcome easily. From this one example, one can see simultaneously the terminological, conceptual and methodological power of information as a concept.

Uncertainty about an outcome is therefore related to the number of possible outcomes of a given operation, information obtained about the given outcome from the possibilities is equal to the extent to which uncertainty about that outcome is reduced. Garner uses the example of tossing a coin. If the coin is two-headed, we know that the outcome will be heads - there is no uncertainty about the fact and therefore there is no information transmitted when heads is the eventual outcome. On the other hand, a normal coin with a heads and a tails, communicates a certain amount of information when it lands, reducing uncertainty about the outcome. The more coin tosses there are, the more uncertainty, and the more information processing necessary to reduce that uncertainty.

Of course, the visual system and the brain more generally typically need to reduce uncertainty about more complex and multidimensional processes than a series of coin tosses, and each of the multiple complex dimensions of a given problem typically bear upon and constrain one another. Each successive coin toss has no bearing on the last - the fact that the last coin came up “tails” has no bearing on the probability of the outcome of the next toss. This is not often the case for “real world” information processes, for which the interpretation of the outcome can often be informed by previous experiences.

What is necessary, then, for a natural information processing system, is a way of keeping a record of previous outcomes - of storing information about processes, outcomes and possibilities that have previously been encountered. In order to facilitate information processing, the system must also be able to call upon information relating to specific entities when those entities are not present in the perceptual field. In short, we have to be able to encode, store and retrieve memories - a concept which also has technical definitions which are not far removed from lay definitions. It is commonly held that memories are not simply chronographic, but relate to specific objects, entities, events and ideas (e.g. Baddeley & Hitch, 2007; Tulving, 1993). Memories relating to such things - which can be called upon to facilitate or constrain online information processing - are known as representations.

1.2.2 Representations

Like other psychological terms, representation has a rich and very vague meaning

Scholink - 1983

In terms of the cognitive psychology of vision, the idea of representations has perhaps best been instantiated by Marr, particularly in his seminal text *Vision* (Marr, 1982). Marr's work provided a series of useful concepts and plausible mechanisms of vision as well as a compelling philosophy through which to understand and explicate the process. He drew a distinction between *what* visual processing takes place within the visual system, and *how* that processing is accomplished - positing that knowledge about the details of neural implementation should be secondary to understanding the underlying computational bases for that implementation, and that the latter could be understood properly in the absence a complete understanding of the former.

In general terms, Marr (1982) defined representations in terms of formal schemes, with a formal scheme being:

> *A set of symbols with rules for putting them together - no more and no less*
>
> *Marr - 1982*

For example, the way that we tend to represent numbers is with the Arabic formal scheme, a string of symbols from the set (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) are put together and taken apart using the rules of addition, multiplication, subtraction and division – to name a few. By this, a representation is the system by which we make explicit certain objects or types of information, as well as a specification for the rules by which they are made explicit. By processing these symbols and rules (a representational scheme) we can arrive at a "description" of any given number – with a description being the output of using a representation. We can "make explicit" any number or describe any given type of information regarding numbers using the Arabic representation, but of course there are usually other types of representation available in the face of a given problem. Marr (1982) raises the important point that:

> *How information is represented can greatly affect how easy it is to do different things with it. This is evident from our numbers example: It is easy to add, to subtract, and even to multiply if the Arabic or binary representations are used, but it is not at all easy to do these things – especially multiplication – with Roman numerals. This is a key reason why the Roman culture failed to develop mathematics in the way the earlier Arabic cultures had.*
>
> *Ibid.*

Whilst this definition can help us to understand what a representation *does*, it doesn't do much to tell us what a representation *is.* To define a representation in cognitive psychology quite simply, it is an internal information-carrying state which relates to a cognitive or perceptual entity and which affects cognitive or perceptual processing. Although there have been a number of criticisms of representations (e.g. Thelen & Smith, 1994), it has been argued that these criticisms tend to be objections to the specific properties of representations as conceived in mainstream cognitive psychology rather than an argument that they do not exist (Markman & Dietrich, 2000). Markman and Dietrich (2000) claim that virtually everyone in the field:

> *Seems to agree on the existence of some sort of information carrying state internal to a cognitive system as well as on the need for these internal states in cognitive theories.*
>
> *Markman and Dietrich - 2000*

In an attempt to try and re-conceptualise representations to suit traditionalist and anti-representationalist thinkers alike, Markman and Dietrich refer to the concept of a *mediating state* - an enduring, discrete, abstract and rule governed compositional structure (that is, a structure with contents). The basic idea of the mediating state is that it is an internal state of a system which carries information used in the furtherance of the systems goals.

Returning to Marr's system of representations with this updated definition in mind, representations (internal, mediating informational states) are assumed to exist at various stages throughout the visual system, progressing from the most "low-level", concrete and least symbolic parts of the visual process - such as the registration of light on the retina - to the most "high-level", abstract and symbolic parts - such as those involved in visually counting discrete objects (Stevens, 2012). Processes at each discrete stage of an information processing pipeline are thought to feed-forward and feedback into other stages, such that "lower" and "higher" order representations influence one another in order to result in a stable and reliable perceptual experience.

For example, we can describe a house using a representation of geometrical objects. A square with an upwards facing triangle on top can be recognised as a simple schematic of a house. Putting smaller squares or rectangles at certain points on the square, representing windows, enriches the representation - making it more recognisable and representative of a real house. We can recognise this object as representing a house because of internal mediating sources of information. Each of the shapes is made up of lines at certain angles, the properties of which define the shape - for example a square and a rectangle both contain four lines at right angles, and are only differentiated by rules relating to the respective lengths of the sides.

In general terms, the line making up a square is seen as being a lower order representation than the square, and the square is seen as being a lower order representation than the house. The window can be recognised as being an entity in its own right - accompanied by its own representations - but also as an embedded part within the house, as part of the set of symbols which characterises the representation of the house - being constrained by rules as to where it might be located in relation to the roof, for example.

Marr points out, however, that the level of representation of a given visual form is important:

> *Even though one is not restricted to using just one representation system for a given type of information, the choice of which to use is important and cannot be taken lightly. It determines what information is made explicit and hence what is pushed further into the background, and it has a far-reaching effect on the ease and difficulty with which operations may subsequently be carried out on that information*
>
> *Marr – 1982*

This tells us first that an object can be represented at multiple different levels, and that the selection between these levels has an impact on subsequent information processing. Not all of the information about a particular object can be processed at all levels simultaneously, and it is thought that these levels are organised into a structure - a structure commonly thought of as being hierarchical in nature.

1.2.3 Hierarchies

> *The hierarchy of relations, from the molecular structure of carbon compounds to the equilibrium of species and ecological wholes, will perhaps be the leading idea of the future*
>
> *Needham - 1932*

Hierarchies - another concept for which the lay definition is basically functional - seem ubiquitous. They are multileveled systems which are structurally and functionally organised in such a way as to minimise redundancy and increase efficiency through the balance of functional and structural integration and separation. On our computer desktop we have folders and files, businesses and institutions have superordinate and subordinate groups of employees, libraries are arranged first in terms of genre and then by sub-genre - these are all examples of hierarchical structure and organisation. Central to the idea is the notion that there are “higher” and “lower” levels of the hierarchy, with one type of operation typically occurring more often at the top, and another at the bottom - such as Marr’s system with abstraction increasing as we progress “up” the hierarchy.

In cognitive psychology, understanding that the operations of perceptual and cognitive systems rely upon the use of internal states, and that these internal states rely upon the storage and transfer of information within and between representations, it became necessary to approach an understanding of the structure of these systems within which such operations occur. Commonly, such a structure is presumed to be hierarchical in nature, and it is important to understand whether this is the case.

The reality of models of cognition is important for neuropsychological interventions such as the diagnosis, assessment and treatment of disorders of the brain, whilst the development of computational models for simulating cognition should rely on known principles of organisation - such as whether it is hierarchical. According to Cohen (2000), neuroimaging data is not of a fine enough grain to tell us directly whether cognitive systems and functions are hierarchical in nature (although significant progress has been made in the intervening decades, and such direct proof may not for long be out of sight (Peelen & Downing, 2023)). Cohen argues that, in the absence of direct evidence, a combination of behavioural, neuropsychological, ontogenetic and logical evidence presents a compelling case.

Thinking about hierarchies in terms of perception, we can first look at some alternatives. Palmer (1977) laid out the extreme views of perception, with their different answers to the basic questions of what constitute the units of perception, and how are they combined in order to form integrated percepts. On the one hand, structuralists contend that the perception of whole figures amounts to the concatenation of simple “sensory atoms” - the lowest and most elemental perceptual forms. On the other hand, the Gestalt school holds that the perception of whole figures cannot be perceptually divided, and that the properties of a whole cannot be determined from the properties of its parts.

These positions hold perception as atomistic or holistic respectively, but neither of these positions seems wholly satisfactory. Palmer (1977) writes:

> *It is eminently reasonable to believe that the perceptual representation of a square includes the representation of lines as sub-parts. But it is also important to realise that the square has attributes of closedness and area that are not attributes of the component lines.*
>
> *Palmer - 1977*

Palmer suggested that a more realistic model of perceptual representation would rely upon multiple levels of representation within a hierarchically organised system, with structural units being defined both atomistically (as a set of parts) and holistically (as a set of global properties) at each level of the hierarchy. He assumed that capturing all of the structural information of a given percept would rely upon the interaction of many levels of the perceptual hierarchy.

As well as in relation to perceptual representations, some form of hierarchical order has been proposed to explain general cognitive processing (Broadbent, 1977; Luria, 1970) and memory (Craik & Lockhart, 1972). Outside of cognitive psychology, hierarchical organisation has been proposed to be a ubiquitous feature of all complex systems (Koestler, 1964, 1968; Von Bertalanffy, 1968).

> *All complex structures and processes of a relatively stable character display hierarchic organisation, regardless whether we consider galactic systems, living organisms and their activities, or social organisation.*
>
> *Koestler - 1978*

In his latter work, Koestler focused heavily on the concept of hierarchies – expounding a number of general principles, applying them to biology, sociology and psychology. He defended the concept against the common charge of being an “ugly” term, which conveys to people the impression of a rigid or authoritarian structure. This is couched in the idea of hierarchies as ranks on a linear scale or ladder, which is not what Koestler means by the term:

> *Its correct symbol is not a rigid ladder but a living tree - a multi-levelled, stratified, out-branching pattern of organisation, a system branching into sub-systems of a lower order, and so-on; a structure encapsulating sub-structures and so-on; a process activating sub-processes and so-on… it is at the same time a conceptual tool, a way of thinking, an alternative to the linear chaining of events torn from their multidimensionally stratified contexts.*
>
> *Ibid.*

In defining the usefulness of the concept in real-world systems, Koestler relayed a parable invented by Simon (1962, cited in Koestler, 1970) about two watchmakers, Hora and Tempus, who each make watches composed of a thousand parts.

Hora, who we could think of as a structuralist, builds their watch from the ground up, assembling hundreds of pieces, from the first to the last. If they drop or pause their watchmaking, the watch falls apart, and Hora must start again from the beginning. Tempus on the other hand makes sub-assemblies of ten parts each, making ten of these into sub-assemblies of one hundred parts each, and then puts these together to finish the watch. If any of these sub-assemblies is interrupted or destroyed, only that sub-assembly needs to be reconstructed whilst the completed ones remain stable.

Tempus' hierarchical method is more efficient, their watches are also more stable and resilient, as well as amenable to repair and improvement. Quoting Simon (1962), Koestler writes:

> *Complex systems will evolve from simple systems much more rapidly if there are stable intermediate forms than if there are not. The resulting complex forms in the former case will be hierarchic. We have only to turn the argument round to explain the observed predominance of hierarchies among the complex systems Nature presents to us. Among possible complex forms, hierarchies are the ones that had the time to evolve.*
>
> *Simon - 1962*

Hierarchies are argued to permit the rapid development of forms from sub-ordinate structures. Viewed through the lens of atomism and holism in perceptual theory, we can imagine that the perceptual hierarchy assembles information into sub-assemblies (or representations) simultaneously across multiple levels - permitting the rapid emergence of stable multidimensional percepts. The emergent perceptual representation - comprised of sub-representations - would be resilient to perturbation, since changes at any scale in the perceptual input would only affect sub-representations at the level of analysis at which it occurred – a sub-assembly of Tempus' watch.

In this way, if a window on our house is smashed, we can still recognise the house as our own - recognising at the subordinate level of the window that a change has occurred without this affecting our perception of the superordinate whole. Although we may speak of a window as part of a whole house, the idea of parts and wholes are, for Koestler, entirely conceptual. In their exclusive and absolute sense, parts and wholes "do not exist anywhere".

1.2.4 Interim summary

Here, we have looked at information, representations and hierarchies and their relevance to the problems of vision and the visual system. Through the combination of these concepts we can conceive of the visual system as involving the hierarchically organised representation of information. Information is the stuff of representation, and representations are the nodes on each level of a richly interconnected hierarchy. Information is processed and uncertainty is reduced through the parallel operation of numerous levels of representation distributed throughout the hierarchy, ranging from the most concrete to the most abstract - the outcome of which we call vision.

I argue throughout the thesis that visual information is processed hierarchically, that is, each of the various levels of representational complexity of the visual input - ranging from its most atomistic to its most holistic values - are processed in parallel throughout a distributed, hierarchical system of representations. The questions addressed within the work pertain to the relative independence of each of these levels of representation, and whether certain levels of representation dominate the perceptual process.

In this work, I am concerned with the visual processing of faces, arguably the most complex biologically and socially important objects that we routinely encounter. As we will discuss, there are reasons to believe that the processing of socially relevant objects such as faces is qualitatively different to that of other types of objects – to the extent that it has been argued that the perceptual processing of socially-relevant phenomena has driven the evolution of the brain (Dunbar, 1998) and face (Adams, Albohn & Kverga, 2017). Such ideas are central to the burgeoning field of *social vision*, which has emerged to deal with the complexity of the perception of socially relevant stimuli, and which is discussed in the following section.

1.3 Social Vision

1.3.1 History and Context

> *Most researchers now agree that visual perception is more than just receiving an image but is clearly related to understanding an image.*
>
> *Nakayama – 2011 (Adams et al., 2011)*

A key figure in the emergence of the field of social vision is Humphrey (1976), a primate vision scientist. His provocative and intuitive argument is that nature does not tend to "tolerate needless extravagance in the animals on her production lines", that the unnecessary typically goes away, and what emerges only does so when it is necessary. In short, he argues that nature is exceedingly economical - and so our sociability and intellectual creativity must be necessary for our su, it cannot be a bi-product. The reason for the expansion of these faculties is, in Humphrey's opinion, the demands of the social environment.

Humphrey (1976) suggests that the period of "normal science" in which he wrote had little space for measurements of the intelligence involved in social transactions, but hinted that a Kuhnian scientific revolution (Kuhn, 1962) may provide room to better understand the nature of human intelligence through the study of the social. He argued that interactions with social compared to non-social objects is characterised by an inherent uncertainty, instability and mutual exchange – by "conversation" or "transaction" rather than by simple use. In this, Humphrey suggests that the psychological truth of social beings such as ourselves may lie in understanding the social aspect of their nature.

Following on from this, Dunbar (1998) described the influential “Social Brain Hypothesis”, questioning why human and primate brains would need to be so complex and powerful if they were simply for the processing of ecological information - since our “natural” environments are not that much more complex than those of other mammals. Such complexity may be necessitated by the need to “mind-read”, to infer the intentions and internal states of others and in order to detect and engage in deception. Dunbar found a correlation between species’ average neocortical volume and the size of their social networks - with humans having the largest of both.

Going further, Adams, Albohn and Kveraga (2017) suggested that the face, as well as the brain, may have developed to facilitate the communication of social information. They cite evidence from primatology, comparing “Old World” and “New World” monkeys. The former have more visual access to one another - living in treetops - and have a greater capacity for facial expression than the latter, who have less visual access - living lower and amongst foliage. At the extreme end of this distribution are the Platyrrhine “night monkeys”, living nocturnally in trees, who seem to have no capacity to emote with their faces.

Buckner and Krienen (2013) discussed how the major cortical expansion in hominids was in association cortex. This suggests that the integration and interrelation of information across and between the senses, behaviour, and higher-order conceptual areas (including those involved in the processing of social information) was of tremendous importance in our development - to which the metabolic expense of developing and maintaining such cortical tissue is testament. Another major difference between primates and other mammals is in the visual system (Kaas, 2013), with more of the brain being dedicated to vision in humans than for all of the other sensory modalities combined (Johnson & Adams, 2013).

What such thinking and discoveries amounts to in social vision research is the central tenet that vision is not just about seeing, but about understanding - that our visual system shapes and is shaped by the information that it processes, and does not simply pass input into a cognitive system for later processing. Social vision is the cognitive science that has emerged in order to try and elucidate the visual processing of social information, to understand the interplay between basic visual processes and pre-existing representations of others.

Social vision deals specifically with vision; how social information is processed visually; the development, maintenance and updating of social representations; and how the interplay of visual processing and social representations affect behaviour. Social vision is viewed by Nakayama (2011) as the result of an “emerging kinship” between social psychology and vision science – two fields often practiced in different buildings and different departments of the university.

Even within the very specific domain of visual face perception - into which the current work fits - emotional expression perception has historically been viewed through a social psychological lens, and identity recognition through a visual cognition lens. Eye gaze has been studied in each domain, but with emphasis placed on different levels of analysis in each case - with social and developmental psychological investigations focusing on communication, and visual cognition focusing on perceptual mechanics. It has only been through the emergence of social vision that the combined and interactive activity of such dimensions of the face in visual perception has been addressed (Hedgecoth, Strand & Adams, 2023).

Vision scientists have developed a number of powerful techniques and concepts for the measurement of basic visual processes, but it has become clear that these basic visual processes support more than just the analysis of shapes and the recognition of objects (Gibson, 1979; Marr, 1982; Peelen & Downing, 2017). Social psychologists have noted how easily and quickly social categorisation occurs (Allport, 1954; Bargh & Chartrand, 1999; Fiske & Neuberg, 1990), drawing attention to low-level visual processing in the formation of impressions about others rather than simply higher-order cognitive states.

Johnson and Adams (2013) argue that because visual perception seems to occur so early in the processing streams which terminate in higher-order cognitive areas, vision scientists have historically presumed that such processes were the products of feedforward or “bottom-up” mechanisms presumed to be impenetrable to and isolated from higher-order cognitive processes - such as those that social psychologists are interested in, like impression formation or prejudice. Such thinking was perpetuated by the use of stimuli to which higher order cognitions would have no relevance, such as Gabor patches - about which a person may be asked to judge by orientation.

When it comes to the visual awareness of more perceptually and cognitively complex stimuli, such as faces, we are likely to be concerned with more than simple orientation, but may instead - looking at a face - have seemingly immediate access to information about a range of socially relevant dimensions - including sex, age, race, emotionality and intentionality (Macrae & Bodenhausen, 2001; Todorov, 2017). For each value of the above dimensions, we possess relatively stable internal representations - for instance of men, of sadness and of youth. We have more various and complex pre-existing, “top-down” cognitive states such as expectations - which refer to the properties of the human face - than we do for Gabor patches. Not only this, but complex objects such as faces are more relevant to a wider array of potential behaviours that are subserved by the visual system than simpler non-social objects, such as those involved in navigation and social cognition (Peelen & Downing, 2017).

It has been argued that mental models, schema, sets of expectations or representations are necessary heuristic mechanisms that enable us to navigate and effectively respond to the complexities of social and personal life (Johnson-Laird, 1983; Macrae & Bodenhausen, 2001; McClelland, 1995; Piaget, 1985). It has also been shown that high-level information about social stimuli can affect the way that they are perceived, for example Anderson et al. (2011) found that when positive and negative “gossip” was paired with faces, those paired with negative information (e.g. “Threw a chair at his classmate”) were perceived more frequently under the conditions of binocular rivalry.

Findings such as these have helped to break down rigid barriers between the sensory and the semantic, perception and cognition, and between “low” and “high” levels of the perceptual process. The emerging perspective of social vision helps us to understand that what we see affects what we think and what we think affects what we see. Even early visual processes can be affected by prior knowledge just as the formation of knowledge and the development of memories are affected by early visual processes. Perception is bidirectional.

An important debate that has helped to instantiate social vision has come from the desire to determine whether social categorisation - or the activation of social representations - from visual information is automatic. Following the work of Allport (1954), it was assumed for decades that social categorisation was automatic - that mere, brief exposure to a person was sufficient to trigger a cascade of classifications relating to their demographics. Lacking an empirical framework within which to test this idea, it remained basically unchallenged for decades (Quinn & Macrae, 2005), forming the cornerstone of several influential theories of person perception (Brewer, 1988; Fiske & Neuberg, 1990).

Recent experiments which take into account the complex and multifaceted nature of the visual process have shown that this is not always the case, and that categorisation can depend upon availability of attentional resources (Gilbert & Hixon, 1991), processing goals (Macrae et al., 1997) and prejudices of the observer (Lepore & Brown, 1997). In other words, categorisation can be modulated by lower-order or higher-order processes.

Investigating the automaticity of social categorisation has necessitated the employment of techniques from vision science, helping to bridge the gap between the two disciplines. The debate has largely moved into attempts to describe the functional architecture of the visual system and the proposed face processing system through probing the interaction of social dimensions and the values of those dimensions of the face in visual processing (e.g. Atkinson, 2005; Ganel et al., 2000, 2002; Graham & LaBar, 2007; Schweinberger, 1999). Work such as this has led to the development of a dynamic interactive model of person perception (Freeman & Ambady, 2011; Freeman, Stolier & Brooks, 2019).

1.3.2 The Freeman and Ambady model

The Freeman and Ambady (2011) model (“the Freeman model) is a dynamic interactive theory of person construal. The model supposes that person perception is achieved through the operations of a dynamic system, in which the continuous coordination and interaction of lower-level sensory information and higher-order “social cognition” across multiple levels of processing gives rise to stable perceptual states, or “person construals”. The model supposes that perception of others emerges gradually through ongoing cycles of interaction between social categories, stereotypes, high-level cognitive states, and the processing of low-level sensory information relating to the face, voice, and body.

This process operates upon dynamically and probabilistically reconstructed, flexible internal representations of categories and stereotypes – rather than static, symbol-like representational structures. These representations are simultaneously and dynamically affected by the state of activation of other representations, the systems history, internal constraints (such as the relative patterns of excitatory and inhibitory connections between representations) and random noise. In this respect, one category is never completely active and another completely inactive – they are more likely to be co-active to varying degrees, with the respective overall state of activation characterising the state of the system.

A typical instantiation of the model involves four levels, the cue level, the category level, the stereotype level, and the higher-order level (see figure 1.1). Bottom-up, sensory signals about the stimulus enter the system at the cue level and top-down, semantic signals at the higher-order level. At any given time, these operations are in continual and often bidirectional interaction. The outcome of this is the settling of the system into a stable attractor state which is consistent with the overall activation of a number of categories consistent with the perceptual target.

Within each of these levels there are one, or several pools of nodes – with nodes each representing some feature or “micro-hypothesis”. For example, within the category level, there is the Sex pool, within which are the Male and Female nodes. Intra-level connections are typically mutually competitive – since the categories are mutually exclusive. For example, White and Black might be mutually exclusive nodes in the race pool of the category level. A person cannot be both White and Black and thus White and Black are mutually exclusive categories, and are represented as mutually inhibitory nodes.

The example of two nodes in the same pool which may have excitatory connections provided by the authors is Aggressive and Dangerous – two categories which are more likely to co-present than Aggressive and Nice. The way that nodes function is in line with theory about recurrent connectionist interactive activation and competition models (e.g. McClelland & Rumelhart, 1981). It is thought that prior activation history, the rate of activation decay and the net input of activation into the node from other nodes affect the activation of a node over time. Inputs into nodes are thought to be altered by a normally distributed pattern of random noise – leading to an inherently probabilistic state of activation at any given time.

Category nodes receive input from cue nodes as well as sending feedback to them. They activate stereotype nodes and receive feedback from them as well. The stereotype level contains one pool of mutually exciting and inhibiting nodes. They feed back to categories and receive input from them, doing the same for higher-order nodes. Nodes at the higher-order level correspond to any number of high-level cognitive states – depending on what is being modelled. These might include prejudices, motivations, processing goals or task demands. Such states can be broadly categorised into motivational systems and top-down attentional systems. Person perception emerges out of these dynamics (see figure 1.1).

> *For example, consider the presentation of a male face. Direct stimulation of the Male Cue node will facilitate the Male category, which will inhibit the Female category, which will inhibit the stereotype, Docile, in turn inhibiting the category, Asian, in turn facilitating the category, Black, in turn facilitating the stereotype, Aggressive, which facilitates the category, Male, which facilitates the Male Cues node, and so on and so forth.*
>
> *Freeman and Ambady - 2011*

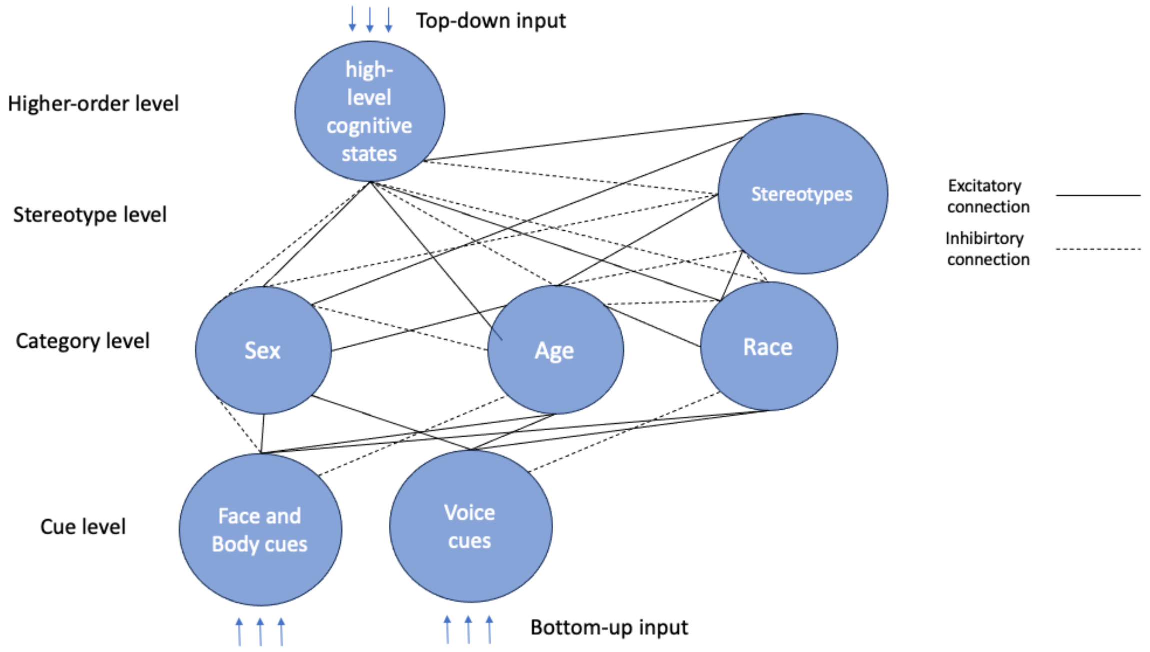


Figure 1.1: Sketch of the basic operations of the Freeman and Ambady (2011) model, showing bidirectional excitatory and inhibitory connections between each level of the network.

The authors suggest that their model makes the novel prediction that an overlap in stereotypical or phenotypical content between two categories could cause the system to bring two categories into interaction – with the former creating top-down pressure on the interaction between categories and the latter creating bottom-up pressure. Of course, some combination of these two types of overlap could bring categories into interaction – and it may be that one type of pressure gives rise to the other through the continuous and dynamic interactions within the system.

In summary, the Freeman model suggests that our perceptions of others emerge probabilistically over time through the continual feedback within and between levels of a hierarchically organised dynamical, recurrent connectionist system. Percepts arise through the continual and mutual effects of bottom-up and top-down signals across multiple sensory modalities. Interaction between perceptual categories can occur as a function of shared bottom-up and top-down properties. In this respect, the model bears the signature of the social vision approach.

Work which takes this approach can be helpful in developing and testing the tenets of extant neurocognitive models of perception, including those which deal with the processing of faces (Bartlett, Abdi & Searcy, 2003; Bruce & Young, 1986; Duchaine & Yovel, 2015; Haxby et al., 2000, 2002). In the current work, I use a social vision approach to develop our understanding of human face processing - comparing my findings with the postulations of existing models and suggesting alternatives or directions for future research. Before comparing these models, I will briefly explore what is known about the process of recognising and processing objects and faces.

1.4 Visual object and face processing

Vision is multifaceted in terms of its function and structure. It can be involved in separating friend from foe, interpreting facial expressions, making social categorisation judgements and identifying threat. This is reflected in the dynamic interplay of various processes occurring simultaneously across multiple distributed networks throughout the brain. When it comes to human faces, we are quickly and accurately able to recognise not only “what” a face is, but “who” the face belongs to (Wallis, 2013), making almost instantaneous judgements about a number of socially relevant phenomena including sex, age, race, emotionality and intentionality (Macrae & Bodenhausen, 2001; Todorov, 2017) - despite distraction from other objects and faces as well as changes in posture, angle, viewpoint and expression.

There are converging lines of evidence from developmental, neuropsychological, behavioural and neuroscientific sources to suggest that we have evolved a specialised and separable functional and structural system for face processing (e.g. Kanwisher et al., 1997; Tsao & Livingstone, 2008). Others argue that such a system is not specialised for faces, but it is merely our ontogenetic expertise with faces which makes their processing appear to be specialised (e.g. Bukach, Guathier & Tarr, 2006; Gauthier & Logothetis, 2000; Gauthier & Tarr, 2002).

Whilst debate in this area continues, there is little argument that faces are not special in some ways (Grill-Spector et al., 2016, 2018; McKone, Kanwisher & Duchaine, 2007; Wallis, 2013). What seems clear is that faces represent a complex, important and commonly encountered type of object that is represented across a diverse range of areas in the brain. Learning about the functional and structural representation of faces in the mind and brain is therefore useful for understanding broader issues in visual object and information processing as well as social vision.

Neuroscientific research over the past few decades has elucidated a network of cortical regions attuned to the job of processing faces (for a review see Duchaine & Yovel, 2015), an understanding of which has become a model for understanding object processing in general - which is one factor contributing to how widely and deeply face processing has been studied (Freiwald, 2020).

Freiwald (2020) laid out four organisational dynamics of the face processing system. First, the concentration of face (or another specific class of object) sensitive cells into object-class specific processing areas; second is the organisation of these areas into a large-scale spatial pattern which is highly conserved across individuals and even between Macaques and humans (Duchaine & Yovel, 2015; Freiwald, 2020). The third principle is the organisation of face areas along functional gradients. From posterior to anterior locations, face-representations become more identity-specific; from ventral to dorsal locations, they become more motion-selective – although it has been argued that motion sensitivity for faces and bodies is subserved by a separate system running from the early visual cortex to the superior temporal sulcus (STS) (Pitcher & Ungerleider, 2021). Finally – according to Freiwald - these cells, areas and gradients are selectively interconnected, forming a network.

From these four principles, we see that there are face selective cells which organise themselves into distinct areas, each seeming to be specialised for a different aspect of face processing. These areas are arranged into a pattern which is conserved across individuals and species, a pattern comprised of separable pathways, or “streams” - which in turn serve distinguishable purposes. These specialised streams comprised of specialised areas comprised of specialised cells form a specialised network - or system. This clear organisation permits investigation into the neuroanatomical basis for theories of cognition, allowing one to map the “hardware” to the “software” in a way that wasn’t possible in the past (Young & Bruce, 2011).

1.4.1 The neuroanatomy of face processing

The cortical starting point for visual processing is thought at the area known as V1*,* in the extrastriate cortex - the first visual area in the calcarine sulcus (Benson et al., 2012), or the “lowest” or “earliest” part of the visual processing system. Signals progress from there through areas V2, V3 and V4 and terminate in high-level visual areas of the ventral temporal cortex (VTC) (Grill-Spector et al., 2018). Between them, the “lower” areas (V1 - V4) are thought to represent simple spatial and featural components of an image, such as the detection of edges in V1 (Hubel & Wiesel, 1962), angled junctions at V2 (Ito & Komatsu, 2004) and shape outlines at V4 (Pasupathy & Connor, 2002). These “low” or “early” visual areas are retinotopically organised - that is, there is a spatial correspondence between cells in the retina and cells in the early visual cortex (Hubel & Wiesel, 1962). The size of the receptive field - the area in space covered by a specific neuron - increases dramatically from V1 through to the “higher” areas of the VTC (see figure 1.2).

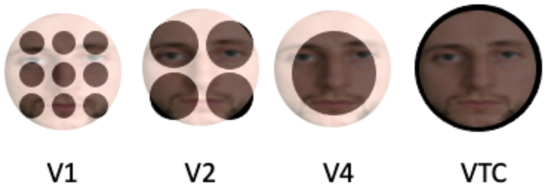


Figure 1.2: Illustration of different sizes of receptive fields, increasing in size with progression throughout the ventral processing stream

The higher areas of the VTC considered "core" to face processing (Grill-Spector et al., 2016) include the occipital face area (OFA) (Pitcher, Walsh, Yovel & Duchaine, 2007), which is thought to detect basic facial features. The OFA may act as an input to both the fusiform face area (FFA) (Kanwisher et al., 1997), which is thought to be implicated in face identity and invariant aspects of the face, and the posterior superior temporal sulcus (pSTS) (O'Toole, Roarke & Abdi, 2002) which is thought to be involved in the processing of dynamic information about faces – such as expression and gaze direction as well as the extraction of form information from motion, enhancing recognition and categorisation processes (O'Toole et al., 2002). Low-level visual areas compute simple, local aspects of visual stimuli such as contrast and orientation, whilst high-levels compute global shape and are implicated in recognition and categorisation (Grill-Spector & Weiner, 2014).

There are three processing streams thought to be involved in face processing, the dorsal "where" or "how" stream, and the ventral "what" stream (Goodale & Milner, 1992) and a third stream which projects from the early visual areas to the STS and is thought to be involved in the processing of dynamic aspects of the face and body (Pitcher & Ungerleider, 2021). Whilst recent thinking about the face processing system embeds it within the visual system more generally, and suggests that visual regions outside of the face system interact meaningfully with face processing, these visual streams – and particularly the ventral stream – are most often studied when considering the neuroanatomy of face processing. Along this ventral stream there is not only the FFA, but areas that appear to be selective for objects such as bodies (Downing et al., 2001) and words (Baker et al., 2007).

Classical models of visual perception were strictly feed-forwards in nature, supposing that signals progressed from low- to high-levels of the hierarchy in a linear sequence. More recent thinking suggests that low-level areas are subject to feedback from higher levels, including information about the perceived global context and task-demands (Lee, 2003). Early stages in visual processing previously thought to exert a solely bottom-up influence over subsequent regions of the visual processing hierarchy (e.g. Pasupathy, 2006; Serre, Oliva & Poggio, 2007) have been shown to be affected by spatial, feature and object based attention (see Cavanagh et al., 2023 for a review). As well as attention, high-level visual representations and memory have also been shown to affect low-level visual representations (Gilbert & Li, 2013; Summerfield & Egner, 2009).

Lerner et al. (2001) found evidence using fMRI of a hierarchical system of representations in the ventral stream which is arranged in bands of increasing sensitivity to image-scrambling. “Higher” or more anterior regions were more sensitive to scrambling – with images being broken into more parts resulting in greater decreases in activation. Their findings suggest that representation goes from local to global, from atomistic to holistic representation as a signal progresses from the lower to the higher areas of the visual processing hierarchy. In another fMRI experiment, Kay and Yeatman (2017) found that responses in high-level visual areas (the FFA and visual word form area) depend upon both low-level visual features and task demands. They suggest that responses in VTC can be computed as the degree to which low-level stimulus properties match top-down category templates.

Taking these two studies together, we can imagine a system of hierarchically organised representations which involve the interplay of both feedforward, bottom-up, stimulus driven information and feedback, top-down, semantic information – arranged from the local and atomistic through to the global and holistic. Considering what we know about the face processing system, this hierarchically organised system involves a division of labour between the properties of the stimulus – such as its variant and invariant or static and dynamic attributes – across specialised cortical areas (made up of similarly attuned cells) and across processing streams.

Perhaps surprisingly, given the long-standing interest in face processing in psychology, there remain some open questions as to how the dynamics described above transform an image of a human face into our perceptions of it. There are a number of currently influential models of face processing which are reviewed below, and which vary in terms of their reliance on neuroanatomical, neuroscientific and behavioural evidence. What each of these models have in common is the idea that face processing is dual-route in nature, involving the operation of two parallel and distinct types of process

1.5 Models of face processing

The following section describes and analyses some influential contemporary models of human face processing. Each of the models described are “dual-route” in nature, that is, they suppose that face perception involves processing along two distinct and separable functional pathways. “The Bruce and Young model” (Bruce & Young, 1986) argues that these pathways are engaged under different circumstances – when processing either familiar, or unfamiliar faces. “The Haxby model” (Haxby, Hoffman & Gobbini, 2000, 2002) – and it’s later revision (Duchaine & Yovel, 2015) – suggests that variant and invariant aspects (or dimensions) of the face are processed down different processing routes in parallel. “The Bartlett model” (Bartlett, Abdi & Searcy, 2003) also involves parallel processing, but argues that the parallel routes are involved in configural and componential processing.

These models differ in terms of their emphasis on face recognition and more general face perception, and their reliance on neuroanatomical and neuropsychological bases. Where they differ most broadly is in where they place the distinction between the two-routes that they stipulate. Experiments such as those reported within the current work have often investigated two dimensions of the face to determine whether they interact in perceptual processing. Such experiments test the tenets of the face processing models such as those reported below, by the logic that two facial dimensions that are processed down parallel routes should not interact, whilst two processed in the same route should interact (e.g. Adams & Kleck, 2003; Dagovitch & Ganel, 2010; Ganel et al., 2005; Schweinberger & Soukup, 1998).

I have set out to understand the relationship between sex and age processing from the bottom-up, I have not set out to test the tenets of any of the three models described below directly. Despite this. Understanding the models and their unique, overlapping contributions to the multifaceted process of face perception will allow us to explore some of the active issues in the face processing literature – particularly with reference to the distinctions that each model is based around. This process will permit us to better frame and discuss the findings of the current empirical work.

Each model has been variously supported and challenged throughout the years, and empirical work and theoretical debate aimed at understanding the process of face perception is ongoing. A review of the evidence for each model is far outwith the scope of this work; instead, what follows is an exploration of face processing through the lens of the distinctions that each model makes. After describing each model and some of the ideas surrounding them, we will discuss the dimensions of sex and age in their terms.

1.5.1 Bruce and Young – 1986

The basic assumption in Bruce and Young's cognitive model is that face recognition and other aspects of face perception rely upon the processing of distinct types of information about the face. They list seven types of information processing that can be carried out in relation to faces (pictorial, structural, visually derived semantic, identity-specific semantic, name, expression and facial speech). They argue that structural encoding processes provide descriptions which underpin the analysis of facial speech and emotional expression codes, as well as for activating "Face Recognition Units" (FRUs).

They argued that tasks such as lipreading and analysis of facial expression take place for all faces through the processes involved in "structural encoding" and "directed visual analysis", but that the numerous operations subserving *identification* were only available for familiar faces, and required access to FRUs – which are representations of individual faces and are elaborated through frequent exposure. In this, there are processes subserving the processing of idiosyncratic and generic aspects of the face – that is, identity-specific and identity-nonspecific processes respectively.

In terms of familiar faces, Bruce and Young suggest that face recognition may rely on a network of representations for different viewpoints, with some representations being attuned to global configurations and others to specific features – in line with evidence from single cell recordings of face sensitive cells in the Macaque infero-temporal cortex that specific cells are preferentially attuned to specific views of the individual (Perrett et al. 1982; Perrett et al., 1984, 1986). Person knowledge is also subserved by access to "Person Identity Nodes", which contain information about names and identity-specific semantic information – such as occupation, friends and where the person is typically encountered.

The activation of a FRU and a PIN engenders the feeling of having recognised a person successfully – identifying their face and knowing who they are, including biographical and personal information about them. This can be achieved through *directed visual processing,* which involves selectively attending to specific visual features as the circumstances demand – such as attending to distinctive features or to pick a friend out of a crowd. Directed visual processing can also be carried out for unfamiliar faces.

The authors make further distinctions between *expression codes,* which underlie the perception of emotional expressions, and *facial speech codes,* which underlie the visual processing of speech. They argue that these types of information are not typically necessary for recognising faces, and are therefore not person-specific – they can be activated with regards to familiar or unfamiliar faces. This distinction is key in terms of the model, as it marks a distinction between the functional components of face processing relevant to familiar and unfamiliar faces.

According to the model, *structural encoding* produces a set of descriptions of the viewed face, including view-centred descriptions as well as more abstract descriptions relating to the global configuration and the features of the face. The view-centred descriptions provide the grounds for analysis of *expression* and *facial speech* codes. The more abstract descriptions (the *expression-independent descriptions* relating to global configuration and facial features) form FRUs. Each of these three types of analysis (expression, facial speech and face recognition) subserve different perceptual classification functions. Their interaction in online processing informs us about the emotional state of the perceived individual, what they are saying and who they are.

The processes subserving the processing of familiar and unfamiliar faces are distinguished between, with the former involving the comparison of perceived faces to FRUs, and the latter relying more upon structural encoding and directed visual processing. These two latter processes build up a representation of an individual over time (FRU), leading to feelings of familiarity. In this sense, FRUs serve as heuristic representations, ensuring that the processes of structural encoding and directed visual processing need to be relied upon only until the FRU is activated.

In summary, the model assumes that the processes underlying the recognition of familiar individuals, and the processing of information about unfamiliar individuals, are independent. The authors argue that representations of familiar individuals are formed through repeated exposure, and that a match between the structure of an encountered face and that of a stored representation of a known individual activates a node in the face recognition network which provides access to information about that individual.

An important aspect of this work is the lack of focus on the brain localisation of the proposed types of informational codes and nodes. The authors stress that the model is a functional one, and that the codes they posit are products of the operation of the functional components of face processing – rather than the functional components themselves. In this, they do not attest to the reality of, for instance, a visually derived semantic code stored somewhere in the brain, but rather that access to visually derived semantic information about a given face can be conceptualised in this way.

The Bruce and Young model was arrived at through a converging operations approach, taking cognitive and neuropsychological studies into account. The model deliberately avoided being influenced by neuroanatomical evidence – which was, at the time, “muddled” (Young & Bruce, 2011). According to the authors, it was quite reasonable at the time to posit a cognitive model without a neuroanatomical basis, with the prevalent analogy being that computer hardware didn’t necessarily stipulate what computer software would do.

The authors didn’t think that the neuroanatomical evidence was strong or relevant enough to constitute one of the approaches taken towards the development of the model, but it gave them “real pleasure” to see the anatomical evidence become clearer, and to become a viable converging operation in developing the next heavily influential model of face processing, that of Haxby, Hoffman and Gobbini (2000, 2002).

1.5.2 Haxby, Hoffman & Gobbini – 2000, 2002

Whilst the Bruce and Young model was focused largely on face recognition, Haxby et al. (Haxby, Hoffman & Gobbini, 2000, 2002) focused on social communication – which they argued to be a more primary and common reason to look at a face. They proposed that changeable (variant) information, such as that which underlies facial speech cues, emotional expression and gaze direction, is the primary means by which social communication is carried out. They argue that, given this, the representation of identity must be relatively independent of the representation of changeable aspects of the face.

According to their model, face perception is mediated by a distributed neural system comprised of multiple bilateral regions. They argue for the existence of a *core* system and an *extended* system, with the core system being involved in the visual analysis of the face for the extraction of social information. According to the original model, the core system involves three bilateral regions of the occipitotemporal visual extrastriate cortex, with each area subserving different roles in different types of face perception. These areas are the OFA (Clark et al., 1996) for detecting faces and their features, and for feeding information to the FFA in the lateral fusiform gyrus (Kanwisher et al., 1997) – which is argued to encode identity information; and the pSTS (O'Toole et al., 2002; Puce et al., 1998) which represents changeable aspects of the face such as emotional expression and gaze direction.

The authors cite a body of neuroanatomical and neuropsychological data to support their argument, including their own findings. Hoffman and Haxby (2000) carried out a selective attention experiment under fMRI conditions in which participants were presented with a series of faces, and asked to determine whether the face on the current trial was the same as on the previous trial in terms of either gaze direction (a variant dimension) or identity (an invariant dimension). Paying attention to eye gaze was associated with greater activity in the pSTS, whilst attending to identity was associated with greater activity in the FFA.

The extended system involves the recruitment of neural subsystems concerned with other ongoing cognitive and perceptual tasks in order to process information about the face – for example the amygdala for processing facial emotion (Breiter et al, 1996), the auditory cortex which has been shown to become active during lip-reading (Calvert, 1997) or parietal regions involved in spatial attention which are activated during gaze perception (Hoffman & Haxby, 2000). More abstractly, we might look at another persons' face in order to gauge their interest, or intentions – such tasks are thought to involve the extended system.

1.5.2.1 Duchaine and Yovel – 2015

The Haxby model suggests then that face perception is a process relying on two distinct information processing routes, one specialised for processing dynamic, variant information about the face and the other for processing static, invariant information. Over a decade later, Duchaine and Yovel (2015) suggested an updated model in which the central distinction between aspects of face processing remains intact, but the two separable routes are the ventral and the dorsal processing streams rather than routes from the OFA terminating at the FFA and pSTS respectively.

They acknowledge the contribution the Haxby et al. model has had in directing research on face processing, but outline areas in which emergent evidence has brought its central neural distinction into question. These areas include suggestions that the OFA might not the input mechanism to the FFA and pSTS, that the FFA may contribute to the processing of variant properties of the face and that several new face sensitive areas have been identified in the temporal and frontal lobes.

In the light of such emergent evidence, Duchaine and Yovel (2015) proposed "a revised neural framework for face processing" in which the known face areas can be divided into two separate pathways. The ventral pathway includes the OFA, FFA, and the anterior temporal lobe face area (ATL-FA) (e.g. Axelrod & Yovel, 2013; Yang et al., 2015); the dorsal pathway includes the pSTS face area (pSTS-FA) as well as the anterior superior temporal sulcus face area (aSTS-FA) (Pinsk et al., 2009) and a region at the right inferior temporal junction which has been shown to respond to eyes (Chan & Downing, 2011), and which the authors refer to as the inferior frontal gyrus face area (IFG-FA).

According to their revised model, the ventral stream primarily represents form information, providing the means to represent invariant features such as identity, sex and age – but which also contributes to expression recognition (given the visual form differences underpinning different expressions). In the ventral stream, face representation occurs first in the OFA in a view-specific manner, the OFA – as well as early visual areas – provide input to the fusiform gyrus. The FFA(s) contain more holistically integrated representations which may be mirror symmetric (Kietzmann, Swisher, König & Tong, 2012). It represents information used for the computation of facial identity and the recognition of facial expressions.

The most anterior area of the ventral stream, the ATL-FA receives input from the OFA and FFA, and it is tentatively supposed to be capable of relatively image-invariant representations of facial identity. Representations at this level permit the encoding of other invariant aspects of the face, and are sharpened through recurrent form processing over the course of several hundred milliseconds through reciprocal connections with the OFA and FFA. Their representation of the function of the ventral stream is therefore hierarchical in nature, with the resolution, robustness and resilience to perturbation of the face representation increasing as the visual signal progresses from posterior to anterior regions.

The dorsal stream shows greater responsiveness to the dynamic or variant aspects of the face, allowing the representation of emotional expression, gaze and mouth movements – although it may also process dynamic information relating to identity recognition. The MT+/V5 area of the dorsal stream as well regions of the posterior-STS (p-STS) have been shown to be involved in the processing of biological motion (Grossman et al., 2005). The right p-STS has also been shown to be sensitive to dynamic faces as well as voices (Watson et al. 2014), suggesting that it is involved in multimodal person perception. Taken together, the dorsal stream is supposed to be involved in processing dynamic person information.

Overall, the authors suggest that the ventral stream represents the structure and surface properties of faces in the posterior face-selective areas and matches these with stored representations of familiar faces in the ATL-FA whilst the dorsal stream extracts information about constantly changing aspects of the face for use in online social communication perception. The

Duchaine and Yovel update furnishes the original Haxby model with a more complex and up-to-date picture of the neuroanatomical bases for face perception.

1.5.3 Bartlett, Abdi & Searcy – 2003

A less well established dual-route model of face recognition comes from Bartlett, Abdi and Searcy (2003). The Bartlett model argues that the routes to face processing are "configural" and "featural". Before properly defining these terms, it is important to state that the Bartlett model is one of face recognition rather than more general face processing, and that the application of the model to more general face processing has led to the emergence of an interactive view of configural and featural processing, rather than a strict, parallel and independent dual-route view (e.g. Kimchi & Amishav, 2010).

Debate around configural and featural face processing, or what is sometimes referred to as "holistic" face processing, remains active (e.g. Amishav & Kimchi, 2010; Behrmann, Richler, Avidan, Kimchi, 2014; Ventura et al., 2022; Zhou, Lin & Zhou, 2022), although there have been relatively few investigations into the relation of such types of processing to social categorisation and the interaction between social dimensions (e.g. Cloutier, Mason & Macrae, 2005; Wiese, Schweinberger & Neumann, 2008). The extent to which different visual dimensions of the face involve configural and featural processing therefore remains largely unknown.

To understand this set of terms in the context in which they are used, it is necessary to take a brief diversion into the holistic processing debate. This will take us through some of the evidence that Bartlett et al. (2003) used to support their model. Contrary to my approach in describing the previous models, I will then look briefly at more recent work which explores the basic distinction between configural and featural processing for more general face processing – since it is that with which the current work is more concerned.

1.5.3.1 Configural and Featural processing of faces

One of the key arguments by which face processing is differentiated from other types of object processing is in the notion that faces are processed more "holistically" than other objects (Piepers

& Robbins, 2012; Richler & Gauthier, 2014; Tanaka & Farah, 1993) – that is, they are more likely to be processed as a Gestalt, or perceptual whole (see Wagemans et al. 2012 for a review). The extreme view is that face parts are not explicitly represented and play no role in face processing (Farah, Wilson, Drain & Tanaka, 1998; Tanaka & Farah, 1993). Some theorists use the term "configural" in a similar vein to "holistic" (e.g. McKone, 2008). The two terms have been argued to be synonymous (Fifić & Townsend, 2010), but other researchers have differentiated the terms (e.g. Piepers & Robbins, 2012; Richler & Gauthier, 2014).

For Richler and Gauthier (2014), configural processing refers to sensitivity to spatial relations between parts of the object (for example, the distance between the eyes of a face) whilst holistic processing is the tendency to process all the parts together as a whole. Maurer, Le Grand & Mondloch (2002) on the other hand distinguish three types of configural processing: sensitivity to first-order relations (such as the mouth below the nose); holistic processing (combining the facial features into a gestalt); and sensitivity to second-order relations (e.g. the distance between the eyes or from the nose to the mouth). They point out that some researchers use the term "configural processing" to refer to one, or all three of these types, and that there is no consensus in terminology.

My preferred definition of the term revolves around the third type defined by Maurer et al. (2002) – involving second-order relational processing. Not only do I see it as the configuration *between* features, but the configuration *of* features. For instance, the length of the brow and the width of the jaw – two aspects of the face known to contribute to sex processing (Rhodes, 2009) – seem to me to be configural rather than featural in nature, affecting the overall configuration of the face and the relation of the features to one another. I define configural processing in terms of the spatial relationships within and between aspects of the face.

In the same vein, featural processing may refer to the aspects of the face themselves, or the aspects of those aspects – depending on the level of processing in question. It may be too simplistic to define these terms without reference to levels of processing. For example, a nose is a feature of a face, but a nostril is a feature of a nose. Similarly, the nose plays into the configuration of the face, and the nose itself has a configuration. Featural processing may also include more “global features”, such as skin texture and colour (Rhodes, 2009). These features have also been defined as “surface cues” (e.g. George & Hole, 1998).

Holistic processing, for me, involves the processing of the features as well as the configural information – that is, the whole face, not just the configuration between and of the features. In this respect, it is the highest form of face processing and involves attending to the face as a whole. In that respect, it may be hypothetical. The fact that I am able, and feel obliged, to render my own definitions is testament to how various and vague the definitions of these terms are, despite their being widely considered to be integral aspects of face processing.

Evidence for increased “configural” or “holistic” processing in faces comes from demonstrations that the inversion of a face leads to greater decrements in recognition performance than inversions of other objects (the “inversion effect”) (Yin, 1969); demonstrations that the components of a face (e.g. eyes, nose) are better recognised within the context of the whole face than components of other objects are recognised within their whole context (measured using the “part-to-whole” paradigm) (Tanaka & Farah, 1993); and demonstrations that aligning two half faces of different individuals makes recognition of one individual more difficult than when the two halves are misaligned (the “composite face effect”) (Young, Hellawell & Hay, 1987) – although misaligned halves have been shown to elicit comparable responses to aligned halves when the faces are dynamic (Zhou et al., 2022).

An alternative to the “configural” or “holistic” view has been characterised as “analytic”, “feature-based” (Fifić & Townsend, 2010) or “componential” (Bartlett et al., 2003), by which it is proposed that face perception is built up from the perception of individual features or dimensions. Diamond and Carey (1986) showed for example that the inversion effect is increased in pictures of dogs for dog experts. They argue that the inversion effect emerges when the members of the class (for example, faces or dogs) share a configuration, can be individuated on the basis of second-order relational features, and the observer has significant expertise to exploit such features in establishing recognition. They suggest that expertise leads to an increase in reliance on lower level representations as opposed to high-level holistic representation.

Findings such as these led Bartlett et al. (2003) to argue that the two routes to face recognition were a slow, featural processing route and a fast, configural processing route – both routes are thought to be independent but simultaneously available and active in perception. The latter route is said to be affected by inversion, whilst the former is only affected through inversion by means of the "assignment of directions" (Rock, 1973) – by which the features are typically encountered, and therefore encoded and represented as upright.

More recently, Kimchi and Amishav (Amishav & Kimchi, 2010; Kimchi & Amishav, 2010) have tested this dual-route model with respect to face processing, rather than face recognition. They created facial stimuli in which they could vary the spatial relations between features, and the features themselves whilst minimising differences in their spatial relations. They conducted a discrimination task and a Garner speeded classification task (Garner, 1974).

In the discrimination task Kimchi and Amishav, (2010) found that the type of processing used by participants was typically determined by the discriminability of the components – such that when differences between features were larger participants relied on features to make their judgements, and when differences in the spatial relations between features were larger participants relied on configurations to make their judgements. They found that discrimination is carried out on a single varying component, leading them to argue that facial components are represented in face processing (contrary to the extreme holistic view (e.g. Tanaka & Farah, 1993)) and are processed independently.

They found that differences in configuration could facilitate feature-based discriminations, particularly when featural discriminability was low. The facilitation was greater than would be expected when featural discriminability was high, leading the authors to suggest that configural and featural processing may interact in processing – rather than being carried out independently and in parallel, leading them to reject this aspect of the Bartlett model. This interactivity between the two types of processing was mirrored in their speeded classification study, in which they found a bidirectional failure of selective attention towards configural and featural face information (Amishav & Kimchi, 2010).

The theory and methodology of the Garner speeded classification task are discussed in detail in chapters 2 and 4 respectively, but for now we can say that Amishav and Kimchi (2010) asked participants to attend to either the configural or the featural information in faces whilst keeping the other type of information constant, or varying it. They created four faces using two sets of featural information (eyes, nose and mouth) and two sets of configural information (inter-eye and nose-mouth distances). Participants were asked to discriminate between the identity of two different faces between trials.

The authors found bidirectional interference for upright faces, that is, slower performance when they varied the irrelevant information compared to when they kept it constant. They found asymmetrical interference for inverted faces, such that variations in features adversely affected the processing of configurations but not vice versa. This led them to suggest that featural processing is dominant in inverted face processing – or that inversion affects configural processing more than featural processing. In a follow-up experiment, they found that selective attention to different facial features was successful, supporting their hypothesis that face features are processed independently of one another.

1.6 The processing of sex and age

The Bruce and Young, and the Haxby models described above argue that certain socially relevant aspects (or dimensions) in the face are extracted along noninteracting and independent processing routes, and that social dimensions processed in different routes do not interact with one another in perception. The emergence of the social vision perspective over the past few decades, which has gradually brought together the emphasis on communication and perceptual mechanisms of social psychology and vision science respectively, has challenged these assumptions (Hedgecoth et al., 2023).

The processing of dimensions of the face in parallel makes sense from the computational perspective of reducing unnecessary interference, but as theorists like Dunbar (1998) pointed out, these dimensions interact in the extraction of social information from other people – in making inferences about their intentions and thoughts. It is only through the interaction of multiple social cues that we can make such inferences, and so it may be that the visual system has evolved to efficiently extract meaningful social information which emerges from their combination (Weisbuch & Adams, 2012).

> *Only in combination do social cues perceptually inform the unified representations that guide our impressions of and responses to others.*
>
> *Hedgecoth et al. – 2023*

A combined, interactive view such as that posited by Hedgecoth et al. is similar to that forwarded by Amishav and Kimchi (Amishav & Kimchi, 2010; Kimchi & Amishav, 2010) with regards to configural and featural processing. Such interactive views hold that bringing multiple separable (but not completely independent) processing operations to bear upon a stimulus facilitates processing in a nonlinear way – that is, in a way which leads to more efficient processing than would be expected by simply adding the effects of each process together. Within the social vision perspective, this process has been referred to as the *shared signal hypothesis* (e.g. Adams & Kleck, 2003) by which it is supposed that the congruency of signals across the values of different dimensions (such as anger and direct eye gaze both indicating approach) leads to facilitated processing.

In terms of the Bruce and Young model, there is some neuroimaging and neuropsychological evidence to support a strict dissociation between identity-specific and identity-nonspecific components, for instance the finding of intact social and emotional categorisation in patients displaying impaired identity recognition (Tranel, Damasio & Damasio, 1988); or the fMRI finding that identity seems to preferentially activate the inferior occipital and temporal gyri, whilst emotional expression and eye gaze (with gaze being something that Bruce and Young admit to having overlooked completely (Young & Bruce, 2011)) preferentially activating the superior temporal sulcus. More recent behavioural evidence has, however, challenged the notion of strict functional separability of these dimensions, advocating instead for integrative interaction.

For example, identity and sex (Ganel & Goshen-Gottstein, 2000; Goshen-Gottstein & Ganel, 2002), emotion and identity (Ganel & Goshen-Gottstein, 2004), and age and identity (Dagovitch & Ganel, 2010) have been shown to be processed in an interactive, or integral manner – that they interact with one another in processing. Similarly, in terms of the Haxby model, behavioural experiments have shown interactive, integrated processing between variant and invariant dimensions of the face; such as identity, emotional expression and facial speech (Schweinberger & Soukup, 1998), sex and emotion (Atkinson, Tipples, Burt & Young, 2005; Wang et al., 2017), and emotion and race, sex and age (Karnadewi & Lipp, 2011). These studies used either Garner's speeded classification task (Garner, 1974) or priming paradigms (see Wentura & Rothermund, 2014 for a review).

For example, Karnadewi & Lipp, (2011) carried out three experiments using Garner's speeded classification task (described in detail in section 2.2). In an effort to test the tenets of the Haxby model, they investigated whether irrelevant variations in sex, race and age (invariant cues) interfered with emotion processing (a variant cue); and whether irrelevant variations in emotional expression interfered with sex, race, and age processing. They found that invariant cues affected variant processing, but not vice versa. They found that the same effect persisted when the faces were inverted – apart from in the case of sex. When faces were inverted, sex ceased to interfere with emotion processing.

The authors' explained this in terms of the disruption of sex processing in inverted stimuli owed to disruptions in configural processing – arguing that inversion caused sex perception to be slowed to the point of no longer interfering with emotion processing. Tasks such as these are often used as tests of selective attention, and have proven to be useful tools for investigating the architecture of the face-processing system.

Such experiments have highlighted that, far from involving independent, parallel processing, the supposed routes to face perception often interact – sometimes in an asymmetrical manner, such that one dimension affects processing of the other more so than vice versa (e.g. Amishav & Kimchi, 2010; Karnadewi & Lipp; Liu et al., 2017). Below, I explore sex and age in face processing with respect to the three face processing models described above.

1.6.1 Sex and Age in face processing

Sex, age and race are sometimes referred to as "the big three" social dimensions, and amongst social perception models they typically hold a unique status (e.g. Gandolfo & Downing, 2023; Jones & Fazio, 2010; Ito & Urland, 2003; Kinzler, Shutts & Correll, 2010; Martin et al., 2015; Stolier & Freeman, 2016). They are considered to be amongst the most frequently activated in daily life, such that they have been argued to be difficult if not impossible to ignore (Brewer, 1988; Fiske & Neuberg, 1990). Categorisation accuracy for these dimensions typically approaches ceiling levels (George & Hole, 1995; Martin & Macrae, 2007).

Despite widespread interest in social categorisation and the general agreement that, if not completely automatic, social categorisation is possible at a glance and with minimal effort, it remains unclear exactly how visual social categorisation is achieved (Kramer, Young Day & Burton, 2017). In the selective attention literature, dimensions tend to be compared to one another in pairs, whilst in types of research concerned with the visual nature of social dimensions, they are often considered by themselves (e.g. Bruce et al., 1993; Rhodes, 2009) – meaning that comparing across different studies can be hindered by subtle differences in the use of terms or the application of methodologies.

More attention has been paid to sex (or gender) in selective attention research than age. It is certainly true that, as socially relevant dimensions, the two have seldom appeared as the principal dyad under experimental investigation (but see Fitousi, 2020; Quinn & Macrae, 2005; Wiese et al., 2008, 2012). One reason for this may be that tests of selective attention to facial dimensions are often used to test the tenets of face processing models such as those described above, but sex and age do not separate neatly across the distinction that each model draws between processing routes. Age in particular has been described as difficult to define in terms of the invariant-changeable continuum (Dagovitch & Ganel, 2010).

For example, there is no obvious way to tell whether sex and age should interact in processing according to the Bruce and Young model – since each dimension seems to be involved both familiar and unfamiliar face processing. In the original paper, Bruce and Young had "relatively little to say" about age in particular (George & Hole, 1995). Sex and age are defined in the Bruce and Young model in terms of visually derived semantic codes which the authors hoped would be delineated and better separated in future studies. Bruce later marshalled evidence to suggest that sex was processed independently of identity, but also showed that sex and emotional expression were separable dimensions using the Garner paradigm (Le Gal & Bruce, 2002) – again leaving no obvious place for sex to fit within the Bruce and Young model.

In the Haxby model, we have dimensions like emotional expression and identity at the respective ends of the variant and invariant continuum – something that changes from moment to moment and something that does not change throughout the lifespan. Dynamic or variant dimensions – such as emotional expression or gaze direction – not only change, but are reversible and have a conceptual baseline (neutral in the case of expression and straight forwards in the case of gaze). The social dimension of age falls into an awkward place between variance and invariance in this regard. Facial age does not change from moment to moment, and is in that respect invariant, but it is variant insofar as it changes constantly throughout the lifespan. It is not reversible, going strictly from young to old, and does not have a baseline – unless we count birth.

The relationship between the concepts of sex and gender is a difficult one, but sex – conceptualised here in terms of gonadal and chromosomal dimorphism at birth – does not typically change throughout the lifespan. Sex is, in these terms, an invariant dimension. With regards to a baseline, a “male bias” has been consistently reported in the face processing literature (e.g. DeGutis, Chatterjee, Mercadeo & Nakayama, 2012; Wardle, Paranjape, Taubert & Baker, 2021). This has recently been proposed to emerge from early developmental experience, with more common visual experience with female faces in early life leading to the development of more narrowly tuned detectors for females – with any faces falling outside of that narrow tuning being categorised as male by default (Gandolfo & Downing, 2023).

The Haxby model itself says nothing, or is “mute” when it comes to “the big three” (Fitousi, 2020), although one may infer that sex and age should interact according to this model given that they can be most readily characterised as invariant – rather than dynamic, variant dimensions which communicate socially relevant information in the same way as emotional expression, gaze direction and facial speech. As for the relationship between sex and age in the Bartlett model, the picture seems more complicated and the evidence seems more diffuse – perhaps owed to the relative dearth of ways of exploring configural and featural processing pointed out by Kimchi and Amishav (2010).

1.6.2 Configural and featural processing of sex and age

Although I'm aware of no systematic study of the relative weight of configural and featural processing in the perception of sex and age from human faces, I make the argument that sex perception is typically more dependent on configural processing relative to age, and age perception is typically more dependent on featural processing relative to sex. That being said, I assume that optimal perception of each dimension requires both types of processing, and that these strategies are flexible, permitting one or the other type of perception to be carried out using either process.

1.6.2.1 Sex

Baudouin and Humphreys (2006) found both inversion and composite effects in sex categorisation – two tests thought to disrupt configural processing. Zhao and Hayward (2013) also found that inversion, scrambling and the use of composite faces adversely affected sex processing – finding that this effect occurred for both Chinese and Caucasian faces for Chinese observers, leading them to suggest that the "holistic" properties of sex categorisation are universal and not race specific. Bruce et al. (1993) also found inversion effects – as long as the hair was concealed.

Bruce et al. (1993) found that a combination of 2-D (local, feature-wise configurations such as brow width and global, inter-featural configurations such as the distance between the eyes) and 3-D configural information provided the best cues to sex discrimination (after naturalistic images). This fits with work by Burton, Bruce and Dench (1993) who used discriminant function analyses with a combination of these types of configural information to create a descriptor which approached human levels of sex categorisation – suggesting that highly accurate sex categorisation can be carried out using just configural information. They also found that, when only facial features were available, sex discrimination accuracy dropped to around 70% - significantly further than was found when only 2-D or 3-D configural information was the only source of available information.

Brown and Perrett (1992) showed participants isolated features as well as faces wherein the features had been swapped with a face of the opposite sex. They found that individual features carried information about sex, but that the order of magnitude of information carried by each feature depended on whether it was presented in isolation or as part of a face. Eyes and brows affected judgements most when presented alone, but when embedded in a different face it was the jaw that had the biggest effect. The finding suggests that featural and configural information are important, but that the presence of configural information changes the nature of the featural processing undertaken.

González-Álvarez and Sos-Peña (2022) created 3D models of faces using Facegen (Singular Inversions, 2008), based on scans of real faces, and faces developed from scratch using the software. They presented both sets of faces with and without cues to surface reflectance (colour and skin texture – featural information), leaving only "structural" – or configural – information. They found that, for both sets of faces, accuracy was at ceiling (~98%) when both sources of information were available. This accuracy dropped to around 80% when just configural information was available – leading the authors to infer that around 20% of the information accounting for sex judgements is extracted from featural information.

Taken together, these studies suggest that, although both featural and configural information is necessary for maintaining high levels of accuracy in sex discrimination judgements, configural information is more important – accounting for upwards of 80% of judgements (Burton et al., 1993; González-Álvarez and Sos-Peña, 2022).

1.6.2.2 Age

According to Rhodes (2009), increasing age involves configural changes up until the age of around 20 – with the development of the forehead, nasal bridge and jaw affecting the overall structure of the face. From there, ear and nose cartilage growth continue, making the ears and nose appear larger. The skin undergoes changes in texture over time (Berry & McArthur, 1986), and by middle-age, loss of adipose tissue results in dropping or sagging of the skin. As age advances, more wrinkles appear and there is a higher likelihood of tooth loss, hair loss and a thickening of the eye-brows. This suggests that, after the age of 20, featural changes characterise advancing age.

These features of advanced age particularly affect the appearance of the eyes and the mouth – with the eyes being found by Jones and Smith (1984) to be the facial feature most indicative of age. The eyes are also the region around which wrinkles are most likely to develop, and the loss of adipose fat makes them appear shrunken and more deeply embedded in the face. Another factor affecting the appearance of the face is changes in cardioidal strain. Increased strain is associated with the chin jutting out and the forehead sloping back (Pittenger & Shaw, 1975), however this is a factor which Hole and George (2011) contend is only likely to affect age estimation up to the age of 20.

An issue with investigating configural and featural information is that the transplantation of features in experimental stimuli can lead to global changes in the configural properties of the face. Controlling for this, George and Hole (1998) found that transplanting older features onto a younger face increased age estimations by 40%, while transplanting younger features onto an older face decreased estimations by 33%. In an earlier study, George and Hole (1995) found that removing information about skin texture (a "global feature", Rhodes, 2009) led to an average underestimation of age by 20 years. This suggests that age estimation can be greatly affected by changes to featural information – but that featural information is not completely independent from configural information.

George and Hole (2000) found that face inversion – typically being thought to disrupt configural processing (see Valentine, 1988 for a review)- and/or negation – which is thought to inhibit shape and shading information (Kemp et al., 1998) – did not significantly disrupt age judgements (in line with Karnadewi and Lipp's (2011) finding that inversion did not affect the interference of age cues with emotion processing). Following this up with blurring, and the presentation of just the inner face area, George and Hole found that the worst performance was when only a blurred, negative inner face area was presented.

Interestingly, this condition has cues from the shape of the head, textural information and information about the features degraded. This means that when all that seemed available was configural information about facial features, age estimation suffered. George and Hole (2000) suggested that accurate age estimation could be carried out based on configural information, but only when the whole head was visible – not just the face. This, in line with the above findings, suggests that configurational information is important in age processing, but not as much as featural information.

1.6.2.3 Sex and Age

The picture painted so far of age estimation suggests that a combination of featural and configural information is used in judgements – and there is evidence of an interaction between these two types of processing. That being said, featural information seems to be a more useful source of information for judging the age of faces. For sex, configural information appears to be more important, with information about features alone seeming only to account for a small amount of variance in categorisation.

In a repetition priming study using ERP which compared both dimensions directly, Wiese, Schweinberger and Neumann (2008) had a prime phase and a test phase. The prime phase consisted of two blocks, categorising faces by sex in one block and age in the other. At the test phase, participants saw the same faces again as well as some new faces, and were required to make either a sex or an age judgement. In the congruent condition, participants carried out the same categorisation judgement on the repeated faces as they did at the prime phase, in the incongruent condition, they carried out the opposite task on those faces.

The authors found positive priming effects for each condition for the age categorisation task, but only for the congruent phase of the sex task. This means that, in the sex task, age related information was also processed – since there were priming effects in the incongruent condition of the age task. In an age processing task, sex information was not processed – since there was no priming effects in the incongruent condition of the sex task.

Wiese et al. (2008) suggest that when one attends to sex information, they invariably process age related information, but not vice versa. They argued that age and sex categorisation rely on “different types of diagnostic information in the stimulus”, suggesting that effective age processing could occur in relation to “surface” level cues, whilst sex processing necessitated configural processing. They couched their interpretation in terms of transfer appropriate processing (e.g. Morris, Branford & Franks, 1977) by which processing is thought to be facilitated by an overlap in processing operations between the encoding and retrieval stages of information processing. By this logic, one processes some age information when carrying out sex categorisation but not vice versa because the operations involved in sex processing overlap with those of age processing but not vice versa.

The ERP data from the same experiment corroborated this view. In the N170 time range, in the age task, the authors found negative N170 peaks over the left occipito-temporal regions of interest for both congruent and incongruent trials. For the sex task, they found small but significant negative N170 peaks in the right occipito-temporal regions of interest for congruent trials. Scott and Nelson (2006) found that the N170 response was greater in the left hemisphere for featural changes to faces, whilst it was greater in the right hemisphere for configural changes.

Also consistent with this the positron emission tomography finding that the processing of face parts are associated with greater activity in the left hemisphere than in the right, and vice versa for the processing of whole faces (Rossion et al., 2001). Taken together, this evidence suggests that age and sex processing may rely more heavily upon featural and configural processing respectively, and that these types of processing may be more reliant upon the left and the right hemisphere respectively.

The N170 response has also been shown to be delayed in processing inverted compared to upright faces (e.g. Eimer, 2000; Itier & Taylor, 2002), which is often interpreted as involving the disruption of configural processing. In another study, Wiese et al. (2012) found that sex categorisation was disrupted by inversion more than age processing, and that sex categorisation tasks were associated with greater activity in the right FFA and left OFA. This – in line with the fact that in both experiments (Wiese et al. 2008; Wiese et al., 2012) age categorisation was more efficient than sex categorisation – led the researchers to conclude that sex categorisation requires greater attentional processing resources than age categorisation, and that age categorisation is carried out in a more automatic fashion than sex categorisation.

Considering this evidence, it appears as though sex and age perception relies more upon configural and featural processing respectively, and that they might rely more upon the right and left hemisphere respectively – although both likely rely on each type of processing and each hemisphere to certain extents. This mutually overlapping division of labour suggests interactivity of processing and that the processing of sex information should be resource intensive relative to the processing of age information.

1.7 Summary

In this chapter, I have characterised vision as a process of hierarchically organised, representational information processing. I have pointed out that vision involves the processing of objects, and the special class of objects, human faces – describing the neuroanatomical system specialised for such processing. I have then described the social vision approach to perceptual processing, focusing on the integration of sensory and psychological information to make inferences about the world – and the primacy of its social nature. I described the dynamical systems model of person processing that has emerged out of this perspective.

I then described some models of face processing – each of which revolve around the idea of a dual-route system of parallel, independent processing of different operations or aspects of the face. I have then outlined the place of sex and age – the facial dimensions of interest in the current case – in the terms of each of these face processing models. I concluded by presenting evidence to suggest that sex and age processing differentially rely upon configural and featural processing respectively, suggesting that they rely more heavily upon the processes of the right and left cerebral hemisphere respectively.

The current empirical investigations involve asking participants to attend to faces and make sex or age categorisation judgements about them. Before we can explore those experiments, we must discuss what I mean by attention, and how the various experimental techniques that I have employed can probe different aspects of the process. Attention is another complex, multifaceted and somewhat nebulous term which has been a central problem in cognitive psychology for decades. The following chapter offers a working definition of attention and describes some of the debates and issues that characterise its study.

Chapter 2: Attention

> *Although the term attention is often used as if its meaning were self-evident, it has remained a remarkably elusive concept*
>
> *Kinchla, 1992*

The problem of visual face perception can be viewed as a problem of attention – a multifaceted and sometimes ambiguous term which is perhaps yet to be more eloquently defined than by William James:

> *Everyone knows what attention is. It is taking possession of the mind, in clear and vivid form, of one out of what seems several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies a withdrawal from some things in order to deal effectively with others*
>
> *James, 1890*

Embedded in this definition is a deeply held assumption about attention – that it is selective in nature. This assumption, as well as some others explored below, characterise the contemporary definition and study of attention. The general nature of James' widely accepted and oft quoted definition has, according to Chun, Golomb and Turk-Browne (2010) been an impediment to its empirical study. The field is so broad and deep they argue, that "Ironically, scientists can't use the term "attention" to "select" relevant studies of interest". The problem is so complex that attending to the right aspects of it is a problem in itself.

To examine how face perception can be understood through the lens of attention, I will explore a few of the general theoretical assumptions held about the nature of attention. To adequately define and understand attention, I explore the interrelated component aspects of attention known as selectivity, integrality and modulation. From this, we shall see that visual attention in face perception may involve a number of discrete stages at which relevant stimulus dimensions are selected, and integrated or separated based on information originating from incoming visual information about the physical properties of the stimulus (bottom-up signals), and from our existing representations of such information (top-down signals).

2.1 Selectivity

The sensory world appears to be full of distinguishable objects, ranging in terms of their complexity and dimensionality. It also seems that human beings are inherently limited in terms of their energetic capacity (Barlow, 1972; Clarke & Sokoloff, 1994), thus limiting the nervous system's capacity to process information. Taking these two facts together – the complexity of the world and our biological limitations – it is generally argued that we cannot process all available sensory information at once (Broadbent, 1957; Kinchla, 1992; Desimone & Duncan, 1995; Summerfield & Egner, 2009).

We must process (or "attend to" (Kinchla, 1992)) the world in a limited and selective manner, selecting a limited number of objects or spatial locations for further processing, and selecting a limited number of actions that we can perform in relation to these objects and spaces – such as discrimination judgements (Duncan, 1984). This limited, selective processing is generally referred to as *selective attention* (see Driver, 2001 for a review). In line with this term, attention has been defined as the mechanism for selecting and prioritising representations (Oberauer, 2019), or as the emergent property of neural mechanisms working to resolve competition for visual processing and the instantiation of behaviour (Desimone & Duncan, 1995).

In practical terms, selective attention is what allows us to follow a specific conversation in a crowded bar, read a newspaper whilst construction work is going on outside, or analyse the drums whilst the guitar is playing. These tasks require selectively attending to a particular, relevant source of information whilst filtering out irrelevant sources of information. Selection helps us not only to concentrate on daily tasks, but is imperative for action and survival. We need to be able to prioritise information for processing; in order to prioritise, we must first select. In vision, it seems that only a fraction of the information available to the retina at any given time can be processed and used to inform behaviour (Desimone & Duncan, 1995), we must therefore be visually selective.

Chelazzi et al. (2010) outlined some computational reasons why the brain implements mechanisms of visual selective attention. First, selective attention can be seen as the mechanism which mediates the choice of the next target for preferential foveal analysis – assisting the oculomotor system to optimise sensory sampling of the environment in line with goals and priorities. Second, selective attention allows us to focus processing on a specific object, permitting us to plan and carry out sequential behavioural processes on that object without being distracted by others. Third, selective attention gates access to perceptual representation to memory systems – that is, it restricts the activation of stored representations to those relevant to the attended object. Lastly, selective attention permits the attended object to be represented in working memory and conscious awareness.

In other words, selective attention lets us focus on one object to the exclusion of other objects, allowing the activation of a representation of that object. This in turn allows us to recognise the object and act upon it meaningfully using stored information about it whilst preventing the representations of other possible objects from interfering with that process. Selectively attending allows us to plan and execute behaviourally relevant actions with regards to specific visual contents whilst minimising interference from others. Some influential accounts suggest that visual objects are in competition with one another for representation, and that this competition is biased towards information that is relevant to behaviour (Desimone & Duncan, 1995; Freeman & Ambady, 2011).

To understand selectivity or competition between representations, it is useful to discuss a long-standing debate within the literature which pertains to spatiotemporal “stages” within the attentional process at which selection is thought to occur. This debate has persisted in some form from the middle of the twentieth century (see Driver, 2001 and Serences & Kastner, 2014 for reviews), and understanding it helps us to understand selectivity and the related attentional mechanisms that we shall come to discuss within the hierarchical context that they are commonly conceived.

2.1.1 The stages of selective attention

The proposed existence of stages in attentional processing date back to the models of Broadbent (1958), Neisser (1967) and Treisman and Gelade (1980). Stages have typically been characterised in terms of the “early” or “pre-attentive” stage, and the “late” or “attentive” stage (see Driver, 2001 and Serences & Kastner, 2014 for reviews). Although there is broad disagreement as to the precise nature of these stages and whether or to what extent selection takes place at each stage, most contemporary models of visual attention appeal to the basic architectural dichotomy between them, and are dual-route in nature (Itti & Koch, 2001; Li, 2002; Treue, 2003; Wolfe, 1994).

Although models differ in terms of how much selectivity they assume occurs at each stage, or whether there are gradual (White, Ratcliff & Starns, 2011) or discrete (Hübner, Steinhauser & Lehle, 2010) changes in selectivity between the two stages, the basic idea remains intact. It is supposed that there is a pre-attentive stage that processes information across large portions of the visual field (e.g. Treisman & Gelade, 1980) or across all visible objects (e.g. Duncan, 1984) in parallel. Simple visual features are assumed to be processed at this stage such as colour, orientation, motion and depth. At the second stage, which is typically supposed to have a more limited capacity and thus be more selective, more complex operations such as face recognition, reading and object identification are thought to occur over smaller portions of the visual field (Wolfe, 1994).

Processing at this stage is typically thought to involve serial rather than parallel processing. In feature integration theory (FIT), processing at this stage involves the binding or integration of features, whereas at the pre-attentive stage only single, separable features can be detected (Treisman & Gelade, 1980). Treisman (1983) argued that it is difficult to imagine attentive processing occurring in parallel across the visual field since conducting operations such as object recognition at multiple spatial locations simultaneously would lead to a combinatorial explosion of necessary computational resources.

Due to the recognition of our computational capacity, and the realisation that certain hypothetical attentional processes would require immense computational power, research in selective attention tends to begin with the assumption that there are “bottlenecks” in the attentional process at which irrelevant information must be filtered out from the relevant. Disagreement over where such bottlenecks and limited capacities lie in the perceptual stream characterises the difference between *early-* and *late-selection* theories of selective attention, such as Broadbent’s filter theory (Broadbent, 1958) and the theory of Deutsch and Deutsch (1963) respectively.

Broadbent's filter theory holds that the bottleneck (or capacity-limit) exists at the sensory processing stage of perception – that we can only sense a certain amount and that irrelevant information is filtered out at the sensory level, before being processed semantically. Later augmentations to this theory (Treisman, 1960, 1969) argue that some irrelevant information may still be subject to sensory processing, but in a highly attenuated way which typically – but not always – prohibits the extraction of semantic information.

According to late-selection theories, sensation is unlimited, but limitations apply to the consolidation of memory or preparation for motor action or behavioural responses (Deutsch & Deutsch, 1963). Unattended stimuli might be subject to full perceptual processing, but the individual may not be able to base their decisions or responses upon that processing or form explicit memories about it. Perceptual processing occurs across the sensory field in parallel according to this view, whilst the capacity-limited, serial stage occurs at the level of selection for awareness, responses and/or the formation of memories (Duncan, 1980).

These theories, then, are differentiated by where in the perceptual process they expect attentional selection to occur – at the level of sensation or at the level of responses. The deadlock between these schools of thought, which seemed incompatible and yet were equally empirically justifiable, has largely been alleviated by the emergence of perceptual load theory (Lavie, 1995; Lavie & Tsal, 1994). Lavie and Tsal (1994) observed that much of the evidence for late-selection occurred in experiments with “low perceptual load”, where there was perhaps only a single distractor and a single target, and/or a relatively simple task to be performed. Evidence for early-selection tended to be found in the presence of a larger number of distractor and/or target stimuli and/or more complex tasks (i.e. a higher perceptual load). From this, Lavie (1995) found that distractor interference decreased as a function of increasing perceptual load – that the harder the task, the less we experienced interference from distractors.

According to Lavie's view, perceptual processing is limited, but automatic, and so we perceive as much as we are capable of perceiving given our limitations. The depth of processing of distractor stimuli depends upon the depth of processing necessary for the perceptual task at hand. If the task meets or exceeds our perceptual processing capacity, there is no leftover capacity for the processing of distractors. The extent to which the distractors are processed is the extent to which there is leftover capacity for perceptual processing given the task.

Selective attention research, then, has moved beyond the simple binary between early- and late-selection and has instead embraced the more complex and continuous problem of the extent to which attention is selective and under which circumstances selection takes place. Selection in visual attention is now considered to operate in parallel throughout a number of levels in the processing hierarchy – within and outwith classically defined visual cortices – with the anatomical and temporal locus of selection depending upon perceptual and cognitive demands as well as by the observers' behavioural goals (Serences & Kastner, 2014).

The empirical work reported below involves the selection of certain facial dimensions and not others, and so we can ask questions about the relative processing demands when selecting one dimension compared to the other, and how this affects the level of the perceptual system at which selection takes place. In order to investigate such issues, we must first understand how selection of object dimensions occurs in the first place, and how differential perceptual interference can arise as a function of which object dimensions are selected and which are not. To understand this, we must explore the distinction between integrality and separability.

2.2 Integration and separation

The empirical work that I report below is typically and primarily concerned with intra-object selective attention – that is, the selection and processing of dimensions within overlapping spatial regions within the same object (a face). As such, I focus here on what Shalev and Algom (2000) call *structural-informational models* of selective attention, and say little about space-based (e.g. Posner, Snyder & Davidson, 1980; Treisman & Gelade, 1980) and/or object-based models (e.g. Duncan, 1984).

Structural-informational models assume that objects are multidimensional in nature, that they are composed of a number of *dimensions* (e.g. Ashby & Townsend, 1986; Garner, 1962), and that we can classify *dimensional values* – or, simply, values (Maddox, 1992) across each perceivable dimension. For example, an object might be a ball, to which we can attend to the dimension of size or colour and classify into the values of “small” or “red” respectively. Similarly, the object might be a face, we could attend to its sex or age and classify values on those dimensions as “male” and “young” respectively.

Structural-informational models are best characterised by Wendell Garner (Garner, 1962, 1974) – a theorist whose terminology and methodology have become tent-poles in the cognitive psychology of attention (Algom & Fitousi, 2016). Garner’s theory is – according to Shalev and Algom (2000) an analytical theory as opposed to a holistic theory (such as the space- and object-based theories). The analytic aspect of Garnerian theory is, according to Shalev and Algom:

> *In the sense that attention subserves the ubiquitous tendency to deconstruct the stimulus, any stimulus, into its constituent dimensions*
>
> *Shalev and Algom – 2000*

They go on to argue that humans are motivated to base their representations upon the fundamental building blocks – or dimensions – of perception, typically attending to or focusing on variations along single, relevant dimensions. In Garner’s theory, attention is directed to these dimensions of experience rather than to continuous wholes or stimuli. That is, we do not attend to the colour, size and shape of objects or spaces, but to colour, size and shape themselves.

Importantly, the unit of selection is also the unit of interference – such that, in the case of analytic models, it is dimensions that we are selective for and dimensions which interfere with one another in processing. In relation to these theories, the problem of empirical import for the study of attention relates to *how* the different dimensions of the stimulus interfere with one another during perceptual processing. This is the problem of integrality, which is most commonly examined through the relative success or failure of selective attention to the dimensions of an object.

Contrary to naïve assumptions, over half a century of research in cognitive science suggests that that certain stimulus dimensions interfere with the processing of others – such that one cannot selectively attend to the target dimension without variations on an irrelevant dimension affecting performance (see Algom & Fitousi, 2016 for a review). Object dimensions are thought to characterise the structure of the object and represent distinct sources of information about it (Garner, 1974). There are therefore multiple sources of information present with regards to any perceivable multidimensional object which may interfere with one another.

This is the problem of perceptual independence (Ashby & Townsend, 1986). If we take a stimulus with two dimensions, *A* and *B*, these dimensions are said to be perceptually independent if the perception of each is not contingent upon or does not interact with the perception of the other (Garner & Morton, 1969). If two dimensions are perceptually independent, they are said to be *separable*, they can be separated in processing and the processing of one does not entail the processing of the other. If two dimensions are not perceptually independent, they are said to be *integral*, they are integrated in processing and one cannot be processed without some processing of the other occurring (Garner & Morton, 1969).

With respect to the question of selective attention, we can selectively attend to separable dimensions, but we cannot selectively attend to integral dimensions – selective attention is said to fail in these circumstances (e.g. Shalev & Algom 2000; Fitousi, 2020). When two dimensions interfere with one another under experimental conditions, this is often referred to as *Garner interference* (Pomerantz, 1986) or the *Garner effect (*Algom & Fitousi, 2016)*,* whilst the relative level of perceptual independence of stimulus dimensions is often discussed in terms of *integrality*.

Garner found, for instance, that hue and shape are separable, and hue and brightness are integral (Garner, 1974). That is, one can attend to variations in hue or shape whilst ignoring variations on the other dimension without Garner interference, but there is a Garner effect when trying to do the same with hue and brightness. Garner and Morton (1969) supposed that separable dimensions combined so as to be measurable according to a Euclidean metric whilst integral dimensions combined so as to conform to a city-block (or Manhattan) metric (see figure 2.1).

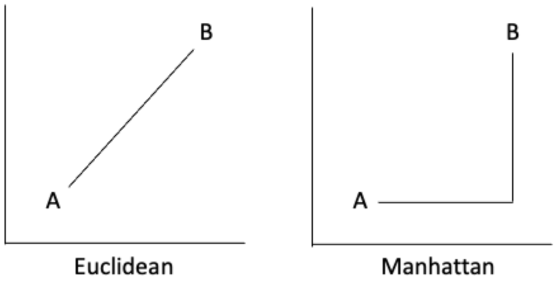


Figure 2.1. Euclidean vs Manhattan metrics of multidimensional stimulus processing. By the Euclidean metric the dimensions (A and B) are directly connected, and cannot be processed independently. By the latter, the dimensions are connected orthogonally and can be processed independently.

It is worth drawing attention to how ubiquitous the distinction between separability and integrality is in psychology. Garner's terminology has been the most influential and broadly adopted in cognitive psychology, but various researchers have described similar processes in different terms. Shepard (1964) and Lockhead (1966), distinguished between unitary and analysable, and nonintegral and integral stimulus properties respectively, Koestler (1964) referred to the integrative and self-assertive tendencies of hierarchically organised cognitive and perceptual systems and Piaget (1985) described the integrative and differentiating aspects of perceptual learning mechanisms. That Garner's terminology is typically used is perhaps owed to the powerful experimental techniques that he designed to investigate the distinction which are still being used in perceptual integrality research today (Garner, 1974).

2.2.1 Stroop and Garner

An accessible way to think about integrality is through the Stroop effect (Stroop, 1935; see MacLoed, 1991 for a review). Stroop presented participants with cards upon which were written the words "red", "blue", "green", "brown" and "purple" 100 times. The words on the cards were either written in black (control) or in one of the colours that the words refer to ("red", "blue" etc). It took participants 5.6% longer to read out the coloured words than it did the words written in black - not a significant effect. When they had to read out the colour of the ink in which the words were written (compared to reading out the colour of the ink in which solid colour squares were presented as controls) they took 74% longer - which Stroop (1935) described as a “marked interference effect”. This type of interference effect came to be known as the “Stroop effect”.

There was a significant Stroop effect for naming the colour of words, but not for naming the coloured words. In Garnerian terms, this is evidence of asymmetrical integrality - or the interference of dimension *A* on the processing of dimension *B*, but not vice versa. It is evidence of successful selective attention when naming words but a failure of selective attention when naming colours. The original Stroop effect was taken as the difference in the time taken to process incongruent stimuli (the word “red” written in green, for example) and the time taken to process the control, but it is often mistaken as the difference in the time taken between processing incongruent and congruent stimuli (such as the word “red” written in green compared to the word “red” written in red). This misunderstanding has persisted despite the fact that Stroop did not use congruent stimuli (Macleod, 1991).

Despite this misunderstanding, the difference between congruent and incongruent stimuli has come to be defined as the Stroop effect in the selective attention literature (Algom & Fitousi, 2016), or as the Stroop-like effect (Fitousi, 2020). Clearly, this definition of the Stroop effect requires stimuli for which the dimensional values can be characterised as congruent or incongruent (known as “Stroop stimuli”, Algom, Chajut & Lev, 2004); for all other stimuli we measure the Garner effect.

Garner interference has come to be understood as a “universal yardstick for assessing the selectivity of attention to any stimulus dimension of choice” (Algom & Fitousi, 2016), regardless of whether the dimensions are “Stroop dimensions” (dimensions with obvious congruities between their values) or otherwise. Garner interference is often investigated using Garner’s speeded classification task (Garner, 1974) which is described in detail in chapter 4.

2.2.2 Stroop and the shared signal hypothesis

The shared signal hypothesis was originally put forwards by Adams and Kleck (2003). In their first experiment, they asked participants to indicate whether faces displayed anger or fear; the faces were either gazing directly forwards (i.e. towards the participant) or their gaze was averted. There was a “pure expression” condition in which the emotional expression was clearly recognisable, and a “blended expression” condition in which the expression was an ambiguous blend of fear and anger.

Adams and Kleck (2003) found that angry expressions were categorised faster when gaze direction was direct, and fearful expressions when gaze was averted. They found that gaze direction significantly biased interpretation of the blended expressions also, such that direct gaze increased the probability of categorisation as angry and averted gaze as fearful. In a second experiment, they carried out the same procedure but replaced the expressions of anger and fear with joy and sadness – finding the same result that categorisation of the former was facilitated by direct gaze and the latter by averted gaze.

The findings were interpreted in terms of the congruency between “approach” and “avoidance” signals between the values of the two dimensions. Anger and joy are supposed to be “approach” emotions, and direct gaze indicates approach. Fear and sadness are supposed to be “avoid” emotions, and averted gaze indicates avoidance. The authors’ argued that the perceptual integration across these dimensions might serve an adaptive function – cuing individuals towards approach-avoidance motivational intentions.

This logic, that specific combinations of values across different dimensions can facilitate processing, has since been tested with regards to a number of dimensions and values. These include race and emotion – such that Eurocentric faces were more quickly categorised as angry and Afrocentric faces as fearful (Adams et al., 2022); sex and emotional expression – such that angry expressions were more quickly detected on male faces and happy expressions on female faces (Becker et al., 2007); and sex and race – such that Asian faces were more quickly categorised when female, and female faces when Asian, and Black faces were more quickly categorised as male, and male faces when Black (Johnson, Freeman & Paulker, 2012).

Kloth, Damm, Schweinberger and Wiese (2015) found evidence for a shared signal between the dimensions of sex and age, showing that categorisation of female faces grows more difficult with increasing age, whilst male faces are categorised more easily with increasing age –

suggesting that young female and old male faces are processed more efficiently than old female and young male faces. They proposed that skin texture was the mechanism for this, with rougher textures associated with masculinity and increased age, and smoothness with youth and femininity.

The link between these findings and the Stroop effect are clear, it appears as though there are values across socially-relevant dimensions of the face which have Stroop congruencies, and which are mutually facilitative in nature. The findings suggest that certain values are perceptually integrated across dimensions. Recently, research in perceptual integrality using the Garner speeded classification task has replicated the findings of Johnson et al. (2012) and Kloth et al. (2015). What remains to be fully investigated is where in the perceptual process these shared signals manifest.

The findings described above have typically been noted when there is some phenotypic or stereotypic congruency between the values of these dimensions. For example, Becker et al. (2007) suggested that the shared signal between Happy and Female, and Angry and Male may originate from (potentially unconscious) expectations about the emotional expressivity of male and female faces, whereas the counter-stereotypic findings of Adams et al. (2022) arose from machine learning processes which suggested that the underlying structures of Eurocentric and Afrocentric faces resemble anger and fear respectively. Similarly, Kloth et al. (2015) suggest that the congruencies between sex and age can be explained at the level of sensory cues, rather than at a higher, representational or stereotypic level. Questions such as these can be posed in terms of attentional modulation, asking whether the effects emerge from the interaction of bottom-up (sensory) signals, or top-down (conceptual) signals.

2.3 Attentional Modulation

If we accept that attention is limited and selective, and involve the integration and separation of dimensions and values, we can then begin to ask what modulates these effects - what factors affect selection and integrality? The problem of attentional modulation remains one of the most fundamental issues in visual attention research (Teufel & Nanay, 2017; Theeuwes, 1993, 2010). According to Theeuwes (2018), all prominent models of attentional modulation which describe selective attention do so with reference to the basic dichotomy between top-down and bottom-up processes, sometimes referring to goal-driven and stimulus-driven selection (Egeth & Yantis, 1997) or endogenous and exogenous attention (Carrasco, 2011; Posner, 1978; 1980). In each case, the distinction is between what impinges upon our attention from the internal, psychological world and the external, physical world.

Bottom-up modulation refers to cases when selection is determined by the feature properties present in the environment. Selection is “under the control of stimulation” (Neumann, 1984), and therefore occurs in a passive and automatic way. Bottom-up attention is supposed to bias the observer towards that which is salient, or stands out from the environment (Itti & Koch, 2001). Stimulus-driven sources of salience include contrast and luminance, or the presence of Gestalt grouping principles such as colour, orientation or shared motion (Treisman, 1999). Salient features in the environment are argued to “pop-out”, capturing attention (Theeuwes, 1993) – a process argued to solely be attributable to bottom-up modulation (Theeuwes, 2010).

Top-down attentional modulation is generally defined in somewhat more abstract terms, but is typically thought to be directed from within (endogenously) and is generally considered a slower form of attention than bottom-up attentional modulation. Top-down modulation is the voluntary orientation of attention (Hopfinger, Buonocore & Mangun, 2000), the deliberate selection of a given location, object or dimension relevant to behavioural goals (Katsuki & Constantinidis, 2014), resulting from an act of will (Baluch & Itti, 2011). Evidence for this sort of attention has come from Posner’s endogenous cuing task (Posner, 1980), in which a symbolic cue or instruction indicates the likely location of an upcoming target. Usually, performance is faster when the central cue is a valid cue to the location of the target compared to when it is not – this is taken as evidence of the capacity to voluntarily orient attention towards that location.

The adaptive selectivity that arises from the interaction of bottom-up and top-down processes is captured by Desimone and Duncan’s biased competition model (Desimone & Duncan, 1995) According to this model, at certain stages in the visual processing stream between stimulation and response there is competition for neural and psychological representation – the distributed neural ensemble responsible for the representation of one, and not the other object or object dimension, must emerge at the expense of other potential such ensembles. Salient stimulus properties (bottom-up modulating factors) and processing goals, expectations and psychological states (top-down modulating factors) impinge upon this process and bias the competition for representation. An application of this sort of idea in the social vision literature is in Freeman and Ambady's (2011) model.

2.3.1 The Bayesian brain as a "hierarchical prediction machine"

The concepts of bottom-up (stimulus-driven) and top-down (psychologically-driven) modulation of attention undergird popular contemporary models of perception and action such as predictive coding (e.g. Clark, 2013; Rao & Ballard, 1999) active inference (Friston, 2003) and relevance realisation (e.g. Andersen, Miller & Verveake, 2022; Reidl & Vervaeke, 2022), all of which are predicated on some version of the Bayesian brain hypothesis (see Friston, 2012 for a review). This hypothesis has been traced to Plato (Friston, 2012), but is more commonly associated with Helmholz (1867) (Clark, 2013; Friston, 2012; Gilhead, Trope & Liberman 2019), who depicted perception as a probabilistic, knowledge-driven process of inference making.

Perceptual and cognitive processes are modulated by bottom-up and top-down factors, but the extent to which these separable sources of information are relied upon is "weighted" according to the inferred "precision" of the information (Andersen et al., 2022; Yon & Frith, 2021). This does not just refer to weighing bottom-up against top-down information, but different sources of information within the same source, for example there may be competition between visual and auditory data, or we may have to "weigh up" our expectations of what a person would like against our knowledge of what they already own when choosing a gift for them.

The various ideas that use the Bayesian brain as a basis rest on the assumption that the brain is predictive in nature; that perception, cognition and action are underpinned by a constantly unfolding attempt to match incoming sensory information with expectations or predictions - or bottom-up information with top-down information. Such thinking characterises the perceptual system as a "hierarchical prediction machine" (Clark, 2013), and is in line with the Freeman and Ambady model of person construal (Freeman & Ambady, 2011).

Bottom-up and top-down signals can thus be viewed as operating upon one another simultaneously and in parallel throughout the levels of a distributed hierarchy of perceptual processing – which ranges from the lowest, most sensory to the highest and most semantic levels. I refer to the "sensory-semantic hierarchy" to capture this axis, arguing that representations on the lower levels of the hierarchy of perceptual processing are more sensory (physical) in nature, and that those on the higher levels are more semantic (psychological/conceptual) in nature. I use the term semantic to refer to "meaning" (another vague and poorly defined term) rather than with relation to its linguistic and logical connotations.

2.4 Summary

This chapter has espoused a particular perspective on attention within which the experiments reported below are considered. This perspective takes the principles of selectivity, integrality and modulation together, framing attention as the result of a number of selective information processing operations unfolding across multiple spatial and temporal stages simultaneously, within a multi-levelled hierarchical and generative system. Intra-object attention involves the selection of relevant dimensions of the stimulus, but these dimensions may interact with one another in processing, and the values of these dimensions may interact across different dimensions. The effect of such interactions may result from the physical properties of the dimensions (sensation) or from their psychological representation (semantic). In this, selective attention to stimulus dimensions may be affected by bottom-up or top-down modulation.

Behavioural research into selective attention allows us to determine the stage (sensory, semantic or response) and cause (bottom-up or top-down) which typically characterises perceptual interference in the processing of a given object. Understanding such effects on attention offers clues as to the structure of the object processing system, allowing us to make inferences about the relative position of different representations of dimensions within the functional hierarchy of visual processing. With specific regards to faces, we can use the concepts and techniques developed in the field of selective attention to interrogate the structure of the face processing system.

In the empirical reports below, the boundary conditions around selective attention to the dimensions of facial sex and age are investigated. Specifically, we ask whether sex and age are integral or separable dimensions, the attentional stage at which any integration might occur, and whether such integration is most readily modulated by bottom-up or top-down signals. In this regard, the work joins the tradition of using the concepts and techniques developed in the cognitive psychology of attention to investigate the functional organisation of face perception.

2.5 Outline of the thesis

The current work investigates the perceptual integrality of sex and age perception in human faces. I carried out experiments using a number of well-validated and extensively tested behavioural paradigms from the cognitive psychology tradition. I took a converging operations approach (Garner, Hake & Eriksen, 1956), defined as the utilisation of more than one set of experimental operations aimed at selecting, testing and systematically eliminating alternative hypotheses in order to arrive at robust and roundly justifiable conclusions about an observed phenomenon.

I have opted for a shallow but broad approach, testing the same questions using a number of experimental designs rather carrying out multiple experiments within a single design – thus reducing the risk that the findings could be attributed to the particularities of one experimental procedure or one set of stimuli. Namely, I conducted a Garner speeded classification task, an Eriksen flanker task, a word priming task, and perceptual priming task and a task-switching task, each aimed at examining how the perception of age and sex from the face interact.

The Garner task gives us a baseline measure of the integrality of these dimensions within the object (the face), the flanker task tests whether sex and age information from irrelevant distractors is integrated into target categorisations. The face priming task gives an index of whether irrelevant sensory information affects categorisation from the bottom-up; and the word priming task tells us whether irrelevant semantic information affects categorisation from the top-down. Finally, the task-switching experiment gives a more direct index of how the sensory, semantic and response level information pertinent to the sex and the age categorisation tasks interact in online processing.

I take a social vision approach, recognising the indistinct boundaries between cognition and sensation, and the importance of social compared to non-social stimulation. The work is strictly behavioural, because the relationship between sex and age processing remains under-investigated, and so there remains a need to explore and elucidate this relationship at the behavioural level.

Chapter 3. General Methods

The experiments reported in the chapters below have each been designed to address specific questions pertaining to the perceptual integrality of visual sex and age in unfamiliar human faces. Methodological similarities between these experiments are discussed below. Experiment specific details are to be found in the methods sections pertaining to each experiment.

3.1 Online recruitment through recruitment service

All of the experiments were carried out online using the online recruitment platform Prolific (https://prolific.co) with participants using their own equipment. This was necessitated by the Covid-19 pandemic, but confers certain advantages. For example, there is evidence that using recruitment agencies or recruiting through social media to carry out online behavioural science experiments leads to the recruitment of a more ethnically and socioeconomically diverse sample than through traditional recruitment of university populations, whilst producing results of comparable quality (Casler, Bickel & Hackett, 2013).

There does exist some debate as to the reliability of data acquired online compared to in a lab, and through recruitment via online platforms such as Amazon's Mechanical Turk (MTurk) (https://www.mturk.com) or Prolific compared to traditional methods. Ramsey, Thompson, McKenzie and Rosenbaum (2016) found that participants recruited through MTurk were more likely to read the experimental instructions than an undergraduate population that was tested in a lab or online. They also found no difference in accuracy on an item recognition task between the two populations.

Uittenhove, Jeanneret and Vergauwe (2022) found a slight and, in their opinion, acceptable drop off in data quality for online compared to in-person testing in a working memory task. In online experiments, they found indistinguishable results between experiments where undergraduate students had been recruited using traditional means and where participants had been recruited through Prolific. There was, however, a drop in data quality compared to these two experiments where recruitment had been carried out through MTurk.

Semmelmann and Weigelt (2017) conducted five basic psychophysics experiments including Stroop and Flanker tasks under lab and web conditions, finding another “acceptable” difference in response time and no difference in accuracy rates between conditions. They also found, however, differences between web browsers in terms of the speed of deployment of stimuli – with Chrome being slower than Firefox by around 4.5ms. Taking this evidence together, I argue that recruiting through Prolific to carry out online experiments ought not to result in markedly different data quality compared to student populations recruited by traditional means, and might additionally confer the advantage of a more diverse sample.

I decided to make the experiments shorter but with more participants than would have been the case if I had conducted them using traditional, lab-based methods. Some benefits of this include avoiding excess participant fatigue and attrition, and also collecting data across a broader range of individuals. I compensated for the increased noise in measurements resulting from short sessions by collecting data from more participants. Taking this step, as well as randomly allocating participants to different conditions, is assumed to minimise any potentially confounding issues – such as browser type (Semmelmann & Weigelt, 2017).

The first experiment (Garner; Chapter 4) was created using Jatos (https://jatos.org) and P5 (https://p5js.org) and the remaining four experiments were created using Gorilla (https://gorilla.sc). Recruitment for all experiments used Prolific, and was subject to selection criteria. Participants had to take part using a desktop or laptop computer and they had to be fluent in English (by self-report). These criteria helped to ensure that participants used computer keyboards to take part and constrained the range of screen sizes on which the experiment would be displayed; they also ensured that participants could read and understand the instructions.

The sampling method was set to “balanced” through Prolific, meaning that half of the participants recruited for any given study identified themselves as male and half as female. Participants who did not report their sex and age demographics were still included in the data – as such individuals constituted a minority and there were no means to screen such participants out using Prolific. Participant sex and age were not included as factors in any of the experimental analyses.

3.2 Design and Procedure

Small pilot studies were run for each of the experiments reported. Typically, these would include one participant per condition of the experiment apart from the task-switching experiment (Chapter 7) where, due to a large number of conditions, a subset of conditions was selected for piloting. In each pilot study, items were informally examined to ensure that accuracy did not fall substantially lower for any particular stimulus than the other stimuli in the set. In that event those items were replaced for the main sample.

Informed consent was given online by clicking “agree” after being presented with general experimental information. Participants were informed that they could revoke their consent by closing their browser window, and that incomplete datasets would be taken as a withdrawal of consent and would not be included in the analyses.

Following the establishment of consent participants were shown more specific experimental information. This consisted of a brief explanation of what each trial would entail, instructions to attend and respond to the *task*-relevant social dimension and instructions pertaining to how to respond. No mention was ever made to the participants of the task-irrelevant dimension in order to avoid drawing attention to that dimension artificially. The exception to this was in the task-switching experiment (Chapter 7), in which both dimensions (Sex and Age) were task-relevant for each participant. The instructions and experiment were presented against a grey background. Participants were encouraged to respond as quickly and as accurately as possible and to otherwise attend to the fixation cross in the centre of their screen.

In each experiment – apart from the task-switching experiment - there were two between-subjects factors, stimulus *set* (Facegen, MPI or Lifespan) and experimental *task* (Sex *task* or Age *task*). In the task-switching experiment, participants took part in both *tasks,* whilst *set* remained a between-participants factor. The decision was taken to keep *task* as a between-subjects factor to eliminate the possibility of perceptual interference at the response level – since the participant cannot get responses mixed up between dimensions if they only know the responses relevant to a single dimension. This was also done to reduce the potential for cross-over effects emerging for participants – a differential difficulty in switching between tasks - where such effects were not of direct experimental interest.

In each of the experiments, participants were instructed to respond using the keyboard keys “F” and “J”, apart from in the task-switching experiment for which the keys were “A”, “D”, “J” and “L”. The levels of the task-relevant variable that each of these keys referred to was counterbalanced between participants, giving different response mappings by which to respond. Participants were randomly assigned to their *task*, *set* and response *mapping* using Randomiser and Quota Nodes on Gorilla. In the task-switching experiment they were randomly assigned to their *set*, *block order* and *mapping.*

The use of Randomiser Nodes on Gorilla works by first randomly assigning a participant to any of the available conditions. The Quota Nodes work by accepting participants into a condition until the maximum number of participants is reached. When the maximum number for a given condition is reached, participants are returned to the Randomiser Node to be re-assigned. This process continues until the participant is assigned to a condition with spaces available. This process means that one participant is always rejected once all of the conditions have a full allocation of participants – the single participant who was rejected was always paid anyway. A participants experience would be functionally identical whether they were randomly assigned to a condition with available spaces on the first try or not – the assignment and re-assignment processes all occur at the back-end.

3.4 Stimuli

Many previous studies of face perception have only tested using one set of stimuli – either drawn from face image databases of volunteer models, or else artificially rendered using software such as FaceGen (Singular Inversions, 2008). Some researchers have mixed stimuli from multiple databases (e.g. Johnson et al. 2012; Xie, Flake & Hehman, 2019). In some cases there are apparently conflicting results between studies using similar paradigms but different stimuli, suggesting that something relating to stimuli may be responsible for those differences. An example of relevance to the current work is that of Fitousi (2020) and Quinn and Macrae (2005), who both carried out Garner speeded classification tasks to investigate the integrality of sex and age, but used different stimuli. Fitousi used photographs from either one, or multiple stimulus sets (depending on the specific manipulation), whilst Quinn and Macrae used “graphics files depicting greyscale images of faces”.

In order to account for the possibility that different stimulus sets may lead to different outcomes, I used three stimulus sets (see appendices 2.1 and 2.2 for the Facegen and Lifespan *sets*, note that the MPI *set* cannot be published here): two “photographic” *sets* which were derived from photographs of real people, and one “artificial” *set* which was generated using FaceGen (Version 3.5). One photographic *set* was derived from the Max Planck Institute FACES database (Ebner, Riediger & Lindenberger, 2010), the other from the Lifespan database (Minear & Park, 2004). Henceforth, I refer to these *sets* as the Facegen, MPI and Lifespan *sets* respectively.

A total of 32 faces were selected from each stimulus set, 8 for each of the four combinations of sex and age (see figure 3.1). The faces taken from each photographic set were selected based on their specific demographics in line with the levels of the social dimensions under investigation (for example: Young and Male). Faces were also selected to exclude hair over the forehead or face, and to avoid images in which facial hair, jewellery or make-up might bias a judgement.

The reason that these three databases were used was that each contained sufficiently large subsets of stimuli to meet the above criteria – allowing us to select a sufficiently sized range of faces for each social category and present them in the absence of superficial identifiers (hair, makeup etc.). The faces used in the MPI database were recruited through a Berlin based model agency – with models needing to meet the criterion of having an “average type” of look, without “eye-catching features”, and be Caucasian. Photographs were taken in a studio using an 8.0 megapixel camera under professional lighting. The resolution of the final images was 2835 x 3543 pixels (Ebner et al., 2010).

The faces used in the Lifespan dataset were photographs of paid volunteers sourced from five locations across Ohio and Michigan – two student campuses, a shopping mall and two senior citizen festivals. The photographs were taken at these locations using a 2.0 megapixel compact camera, using the built-in flash module and “natural lighting” – that is to say, without professional lighting. The resolution of the final images was 1760 x 1168 pixels (Minear & Park, 2004).

For the “Lifespan” database, the Young stimuli (both Male and Female) were taken from the subset of participants aged between 18 and 29; the Old stimuli from the subset of participants aged 50-69. For the MPI database, the ages of the faces used were not provided – only the age ratings from participants in a validation study (see https://faces.mpdl.mpg.de/imeji/). I selected a group of faces and then analysed the age ratings. The range of mean age ratings for the young faces (male and female combined) used in the experiment was 23.55 - 32.72, and the range of mean age ratings for the old faces (male and female combined) was 61.36 – 71.82. Although these ranges are for ratings, not for the actual age of the faces, they suggest that the range of ages for the young and old groups are slightly narrower for the MPI compared to the Lifespan *set*. Using Facegen allowed me to create distinct identities which I could then closely match in terms of the position they occupied in terms of the “age” and “sex” options.

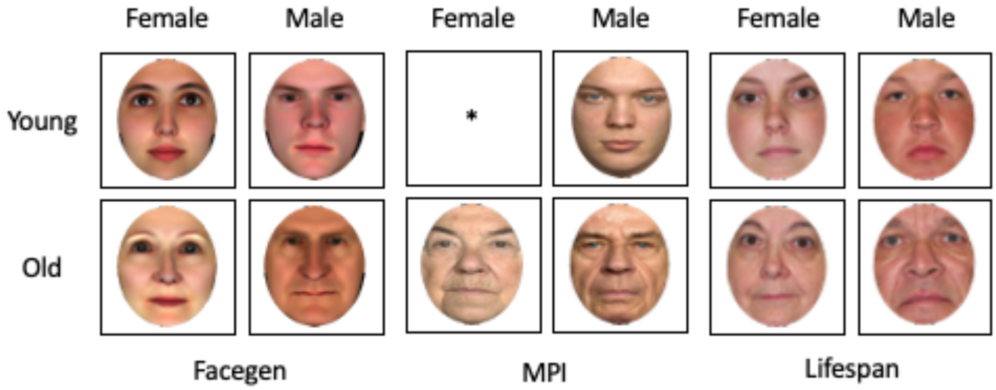


Figure 3.1: Examples of stimuli from each *set* for each level of the two dimensions of interest, Sex and Age. *For the MPI *set*, only the above examples are permitted for publication by the resource holders.

The Facegen *set* was created by setting the Race sliders close to the maximum for European and close to the minimum for the other available races. These sliders were set close to, instead of at, the relevant minima and maxima because setting them there rendered the faces too indistinct from one another, based on informal judgments. For both the sex and the age dimensions there are sliders for shape and colour, which can be locked and manipulated together, or can be manipulated separately. I locked these so as to manipulate them together.

The sex and age sliders (both colour and shape) were set to the relevant levels of those dimensions – close to the minimum and maximum ends of the respective sliders. Identities were then generated by clicking the “generate a random face” button. In some cases, parameters were then manually adjusted to reduce aspects of appearance that might bias a judgement – such as darkness around the upper lip or colour of the lips – and to ensure that stimuli were satisfactorily distinct from one another (based on the informal judgments of the experimenter).

All stimuli were formatted to 400x400 pixels and an ellipse (200x275 pixels) was placed around each face in order to remove any hair and the neck from the images and to normalise them across sets. Hair has been shown to be a heuristic feature for the categorisation of sex (e.g. Goshen-Gottstein & Ganel, 2000) and age (e.g. Rhodes, 2009), and so this was removed so as to avoid it being relied upon to make judgements, and to ensure that judgements were made with respect to the face. Faces were shown in their native colour and texture, even though it seems that face colour varies in general between the sexes (Nestor & Tarr, 2008) as does skin texture across the lifespan (Burt & Perrett, 1995).

The decision to allow features such as colour and texture to vary between images naturally was a conscious one, since I am interested in the processing of sex and age under natural conditions. The variance between faces and *sets* may result in the emergence of differences in task-difficulty within *sets,* but to try and normalise them may result in the removal or degeneration of natural cues to sex or age processing – and so the trade-off of differences in task-difficulty in favour of more naturalistic stimuli was accepted.

All of the stimuli – whether artificially generated or derived from photographs – appeared Caucasian. This decision was taken deliberately. Homogeneity of the race of faces avoids low level noise due to significant differences in lightness and contrast between races affecting the results. At the high-level, Race is of course a highly salient social dimension – often considered one of the “big three” social dimensions along with sex and age (e.g. Jones & Fazio, 2010; Kinzler, Shutts & Correll, 2010; Stolier & Freeman, 2016). As such, race could not covary with sex and age without being accounted for as an experimental variable and systematically varied with the same degree of care in relation to its other covariates.

3.5 Dependent Measures

3.5.1 Inverse Efficiency Scores (IES)

For each participant, in all tasks, inverse efficiency scores (IES) calculated for each condition of interest were used as the primary measure of performance. An IES is calculated by taking the mean reaction time of trials in which the participant was accurate and dividing that value by the proportion of correct responses (PC). IES was proposed by Townsend and Ashby (1978, 1983) as a way to combine RT and proportion of errors (PE), combining speed and accuracy values into a single value, making the data easier to summarise. It can be thought of as an observable measure of the average "energy" consumed by the system over time, or the power of the system. This is based on the assumption that the system has a constant rate of energy consumption, that the perceptual system – in this case – continually uses energy (Townsend & Ashby, 1983), and that this energy is used to the end of efficient processing of information.

Bruyer and Brysbaert (2011) suggested that IES ought only to be used in cases where the proportion of errors is less than 10%. With more errors, there are too few correct responses to estimate the correct latencies and a higher proportion of the data points will be the result of guesses. Another problem with using IES with too low a PC is the fact that the multiplication of RT occurs in a non-linear fashion with decreasing PC (for example, RT is multiplied by 1.1 when PC = .90, but 1.7 when PC = .70). In all chapters that follow - following exclusions for overall poor performance (described below) - mean percent correct was above .95 , thereby supporting the use of IES (see appendix 1.1). As with reaction time data, lower efficiency scores (IES) reflect better performance.

3.5.2 Difference Scores

The primary dependent variable in each analysis was a difference score, generated by subtracting the IES in one condition from another. Difference scores permit us to collapse two variables into one, thereby eliminating a factor from the analyses. This makes the data much easier to visualise and interpret, and reduces the number of comparisons that need to be made. The difference between two conditions of interest constitutes an effect. Specifically, in each chapter below I use difference scores to define a Garner effect, a Stroop effect, a priming effect, and Eriksen effect, and a task-switching effect.

The presence of an overall effect is interpreted as a y-intercept in an ANOVA model that is significantly different from zero ($\alpha \leq 0.05$). A significant y-intercept represents a deviation of the DV from zero independently of any effect of the other variables in the ANOVA. Likewise, a significant main effect of another variable, *task* for example, tells us that the difference score depends on the *task*.

3.6 Exclusions

Prolific (https://prolific.co) automatically sets a default time limit which is five times the expected duration of the experiment (under 12 minutes for each experiment). This expected duration was determined by the experimenters trial-runs of the experiment. Any participant who exceeded the time limit was categorised within Prolific as a “time-out” and their data was not included in the analyses. If a participant terminated their session before completing the experiment, their incomplete dataset was not included in the analysis.

The post-data collection filtering consisted of a few steps. The first criterion for exclusion at the participant level was falling below 2.5 SDs of the mean in terms of accuracy or above in terms of RT, collapsing over all conditions. In each experiment, a second wave of recruitment was then carried out in order to replace the excluded participants. These new participants were subject to the RT and accuracy exclusion thresholds of the *original dataset*. No new data was collected following these exclusions. At this point, individual trials in which response times were longer than 2500ms or shorter than 200ms were removed from analysis.

The second criterion for exclusion at the participant level was scoring above or below 2.5 SDs of the mean in terms of absolute difference scores, again collapsing over all other conditions. Since distinct analyses within the same experiment required the calculation of distinct difference scores, participants were excluded from specific analyses only. For example, a participant might be considered an outlier in the main analysis, but not the Stroop analysis in any given experiment. In such a case, they would be excluded from the former but not the latter analysis.

Excluding on the basis of RT and accuracy separately is a more conservative and more reliable method than excluding based on IES. In terms of the reliability, we can take a hypothetical example of a participant who scored a low mean response time of 400ms, and a PC at the chance level of 0.5. If we calculate their IES, we would get a value around 800 (400/0.5 = 800), roughly the same value as we would get if the participant performed with a more realistic mean response time of 800 with a high accuracy of 98% (800/0.98 = 816). It would actually appear as though the first participant in this example had performed better than the second despite performing at chance levels of accuracy, and would not be considered an outlier in terms of IES. Excluding according to RT and accuracy prior to calculating IES prevents such issues from occurring.

Excluding on the basis of outlying difference scores for individual analyses is also a conservative method which prevents the rejection of participants wholesale in favour of rejection from specific analyses where their results were outliers. This seems important in experiments such as ours which were designed to have fewer trials but more participants. In such experiments – particularly held online, where there is a potential for more latency and connectivity issues – the results are more sensitive to the influence of sequences of trials or even single trials. The current, specific and targeted exclusion criteria permit a conservative approach whilst at the same time retaining as much data as possible.

3.7 Stroop Analysis

Previous literature in the field has suggested that there exists specific congruency or redundancy effects between specific levels of the two social dimensions under scrutiny – namely, participants appear to be more efficient at processing Young Females and Old Males compared to Old Females and Young Males (Kloth et al., 2015, Fitousi 2020). Such a finding has been described in terms of the shared signal hypothesis (e.g. Adams & Kleck, 2003). In order to investigate this effect with multiple stimulus *sets* across experimental *tasks* and a number of experimental paradigms, I conducted a Stroop analysis in each of the experiments.

In each of the five experiments – and for each of the three stimulus sets - I compare performance on “Compatible” (Female/Young or Male/Old) and “Incompatible” (Female/Old or Male/Young) targets. I did this by creating a difference score, subtracting mean IES for Compatible targets from Incompatible targets. A positive difference score would show facilitated performance for stimuli with Compatible compared to Incompatible values across dimensions, and a negative difference score would show facilitated performance for stimuli with Incompatible compared to Compatible values across dimensions. Either type of difference score is taken as evidence of Stroop interference.

This analysis was carried out independently for each of the experimental *tasks* (Sex or Age) without regard for the specific nature of the experiment in question – meaning that I took no account of variables besides *task* and *set* apart from where otherwise stated. This provides us with a large amount of experimental data upon which a common analysis can be performed – allowing us to probe this effect across experimental designs.

Stroop-like tasks are often referred to as “Stroop-like effects” – perhaps owed to the fact that the compatibility between dimensional values in such experiments are not necessarily intuitively obvious from the offset. Within the current work, I refer to “Stroop effects” instead of “Stroop-like effects”, since the effects noted are evidence of interference between compatible and incompatible values across stimulus dimensions.

3.8 Conventions

Throughout the entire thesis, conditions are reported in italics and the levels of those conditions are reported with capitalisation. As an example, I might refer to the Sex *task* or the Mixed *block type*, since “sex” and “mixed” are levels of the conditions “task” and “block type”. The values of the levels of conditions are also capitalised, for instance when referring to “Male” or “Young” when describing these values as relevant to the Sex and Age *tasks* respectively.

Descriptive statistics are ordered according to the above convention and not by order of magnitude. T values are reported in absolute terms and, as well as providing means and standard errors for each level of the relevant conditions, a statement of the direction of the difference is provided. Where means are reported, they are followed by standard errors in parentheses. Significant interactions in ANOVAs are followed up by carrying out a series of independent or paired-samples t-tests for each level of a given factor, unless otherwise stated. Error bars represent 95% confidence intervals in all figures.

CHAPTER 4: Investigating the integrality of sex and age using the Garner speeded classification task

4.1 Abstract

In order to understand whether sex and age – and the respective values of these dimensions – are perceptually integrated in visual processing, I present different combinations of these dimensional values to participants in a speeded classification task. The task measures the Garner effect, the decrement in processing efficiency between conditions in which the task-irrelevant dimension varies compared to when it remains constant between trials. A significant Garner effect is taken as evidence of perceptual integrality between the task-relevant and task-irrelevant dimensions. The full design of this experiment – incorporating redundancy and Stroop analyses – as well as the presentation of different tasks and three different stimulus sets between participants, helps to elucidate how and at what level of the perceptual process interference between these dimensions occurs. I found bidirectional and asymmetrical Garner interference that was homogenous across stimulus sets. This showed that irrelevant variations in sex information adversely affected age categorisation performance to a greater extent than vice versa. I found Stroop effects which showed that young female, and old male faces were processed more efficiently than old female, and young male faces. The task (sex or age) in which the Stroop effect was largest varied between stimulus sets. The pattern of redundancy effects suggests that redundancy between the structure of stimulus values across dimensions facilitates performance relative to the baseline. The modulation of the Stroop effect by stimulus set, and the lack of modulation of the Garner effect by stimulus set, suggests that the integration characterising the Stroop effect occurs at a lower, earlier, visual stage of perceptual processing compared to the Garner effect, which occurs at a higher, later, semantic stage.

4.2 Introduction

Using the tools of cognitive science, we can ask questions about the processes underpinning selective attention to different stimulus dimensions and dimensional values. We can ask about how these dimensions are integrated into perceptual wholes or kept perceptually separate from one another. We can ask whether and to what extent stimulus dimensions interact during processing, and whether their respective processing requirements influence the nature of such interactions (e.g. Ashby & Townsend, 1986; Ganel et al., 2000, 2002, 2003, 2004, 2005; Garner & Morton, 1969; Johnstone & Downing, 2017). The answers to such questions can be revealing about the underlying functional nature of perceptual systems.

Interacting dimensions are known as “integral” and non-interacting dimensions as “separable” (Garner & Morton, 1969). The measure of this interference is known as “Garner interference” (Pomerantz, 1986) or the Garner effect (e.g. Algom & Fitousi, 2016). Integral and separable dimensions can be distinguished experimentally using a number of techniques, including Garner’s speeded classification task (Garner, 1974) which has been argued to be the most precise measure of integrality (Maddox, 1992). It involves participants making decisions about a task-relevant dimension whilst ignoring variations in a task-irrelevant dimension – for example they may be asked to attend to the height of a triangle whilst ignoring the length of one side (Pachella, Somers & Hardzinski, 1981) or attend to the sex of faces whilst ignoring their familiarity (Ganel & Goshen-Gottstein, 2002).

An example of a set of separable dimensions are hue and shape, and an example of a set of integral dimensions are hue and brightness (Garner & Felfoldy, 1970). More complex sets of dimensions have since been investigated, for instance it has been found that the texture and shape of visual scenes are asymmetrically integrated – such that texture affects shape perception more than vice versa (Tharmaratnam, Patel, Lowe & Cant, 2021), it has also been shown that judgements of length are affected by irrelevant variations in duration – such that judgements of length increase when an irrelevant distractor appears for 2.4 seconds compared to 1.5 seconds (Zakay, Bibi & Algom, 2014). Recently, there has been a surge of interest in the integrality of socially-relevant visual dimensions of the human face, for example sex and emotional expression

(Atkinson et al., 2005) or emotion and identity (Ganel & Goshen-Gottstein, 2004) (see Algom & Fitousi, 2016 for a review).

4.2.1 Perceptual integrality of facial dimensions

Ganel et al. (2002) showed that the dimensions of familiarity and sex were strongly integral – that processing of one could not take place independently of processing of the other. In their first experiment, Ganel et al. (2002) employed the typical structure of the speeded classification task. They showed participants familiar and unfamiliar faces which were either male or female. The faces were edited to remove their hair, since hair has been shown to be a heuristic feature for the categorisation of sex (e.g. Goshen-Gottstein & Ganel, 2000). In a within-subjects design, participants were instructed to attend to either the sex or familiarity of faces. There were *baseline* and *filtering* – or *orthogonal* - blocks. In the baseline blocks, only values on the task-relevant dimension varied whilst the task-irrelevant dimension remained constant. In the orthogonal blocks, values on both the relevant and the irrelevant dimensions varied. Garner interference is said to occur when performance is worse in the orthogonal blocks than in the baseline blocks, and indicates that participants cannot selectively attend to the target dimension without processing the orthogonal variation on the irrelevant dimension. Ganel et al. (2002) found Garner interference for both sex and familiarity.

With regards to models of face processing – many of which are based on the idea of parallel-route processing (Bartlett, 2003; Bruce and Young, 1986; Haxby et al., 2000, 2002) - the logic of experiments such as that of Ganel et al. (2002) is that the dimensions processed within the same route should interfere with one another and thus be integral, whilst those processed in separate routes should not interfere and thus be separable. Independence of processing should be mirrored by perceptual independence. By showing that familiarity and sex were integral, Ganel et al. (2002) argued that these dimensions were processed within the same cortical route.

This is a challenge to the Bruce and Young model, by which the processes underpinning familiar and unfamiliar faces should be separable. The results are consistent with the Haxby model, since sex and identity are both invariant dimensions of the face. In terms of the Bartlett model, the picture is less clear, since the precise nature of the reliance on configural and featural information in making sex and identity categorisation judgements remains uncertain (Baudouin & Humphreys, 2006; Burton, Schweinberger, Jenkins & Kauffman, 2015).

Generally speaking, the results of such selective attention experiments concerned with Garner interference have been more complex and have not yielded easily to the simple dichotomy between integral on the one hand and separable on the other. A common finding is an asymmetric pattern of Garner interference between facial dimensions. Graham and LaBar (2007) found that expression judgements were separable from variations in gaze direction but that expression variations interfered with gaze judgements. Because expression judgements were found to be faster than gaze judgements, the authors reasoned that expression judgements were processed before gaze direction could interfere with processing, but not vice versa. To test this, they increased the difficulty of the expression judgements and subsequently found symmetrical interference between the two dimensions - that the pattern of asymmetry disappeared.

This finding - that equalising the discriminability of dimensions diminishes asymmetrical interference patterns - is consistent with the argument of Melara and Mounts (1993), who suggested that information relating to the more discriminable dimension will interfere with that of the less discriminable dimension – and thus that asymmetrical interference would be inevitable when one task was easier, or “more discriminable”, than the other. Whilst this *speed-of-processing account* (SoPA) may explain the results of Graham and LaBar (2007), it has been found by some researchers that asymmetrical task difficulty cannot fully account for asymmetrical patterns of integrality.

For example, Atkinson et al. (2005), showed asymmetrical Garner effects between sex and emotional expression, such that variations in the former would interfere with the latter but not vice versa. They also found that sex judgements were faster than expression judgements. When they equated the task-difficulty (by using different degrees of morphing between values on the task-relevant dimension), the asymmetric pattern of integrality persisted. Schweinberger et al. (1999) observed a similar finding, that the asymmetric pattern of Garner interference between identity and expression held even when identity judgements were made less discriminable than expression judgements through stimulus morphing.

Atkinson et al. (2005) interpreted their results in terms of computational stability, arguing that perceptual representations of variant properties of the face (such as emotional expression) are inherently less stable than those of invariant properties because of their constant capacity for change. Variant properties change, and our representation of the values of those dimensions must therefore be subject to change – or more flexible. They reasoned that variant dimensions are more context sensitive than invariant properties of the face. In this case, expression is more contextually dependent on sex than sex is on expression.

Evidence for this includes that Atkinson et al. also found that participants were slightly slower to judge a fearful face as male than a happy face. They argue that this is consistent with the general finding that women are typically more facially expressive of both positive and negative emotion (Kring & Gordon, 1998), though men are more facially expressive of anger (Brody & Hall, 2008). Atkinson et al. argue that our expectations about such facts make our visual emotion processing systems sensitive to differences in sex, such that we may be less sensitive to male displays of negative emotion and more sensitive to their displays of anger.

Support for this idea comes from Becker et al. (2007), who found perceptual connections – or a shared signal (see Adams & Kleck, 2003; Hedgecoth et al., 2023) – across a number of judgement tasks between anger and maleness on the one hand, and happiness and femaleness on the other. Taking their findings together, Atkinson et al. argued – in support of the Haxby model – that invariant dimensions serve as more stable perceptual referents to variant dimensions than vice versa.

Schweinberger et al. (1999) reasoned along similar lines, suggesting that identity is a more "basic" type of information about the face than expression. They argued that emotional expressivity varies idiosyncratically – as a function of identity – and that taking identity into account might provide a reference to which more variant changes may be more easily processed. Replicating Schweinberger et al. (1999), Ganel et al. (2004) found unidirectional Garner interference patterns showing that identity information affected expression judgements but not vice versa.

They went on to find, however, that when faces were familiar as opposed to novel, Garner interference became bi-directional. The interference was still asymmetric (greater for identity on emotion than vice versa), but the familiarity of the faces was a factor, leading to greater interference of emotion on identity. Like Schweinberger et al. (1999), Ganel et al. (2004) also found that identity was more discriminable than emotion, but found the same pattern of integrality in a follow-up experiment in which they reversed the discriminability of the dimensions. They proposed the *structural-reference hypothesis* as an interpretation for this effect.

4.2.2 The structural-reference hypothesis

According to the structural-reference hypothesis – and in conjunction with Schweinberger et al. (1999) – Ganel et al., (2004) argue that expressions can be seen as variations in the structure of a face (i.e. its identity). As such, each individual will express different emotions in an idiosyncratic way. They propose that knowledge of the underlying structure of a face can be used by the perceiver to process emotional expressions more efficiently. Bolstering this hypothesis, they found that expression judgements were faster for familiar faces compared to unfamiliar faces in baseline blocks, suggesting that the identity of the face – though irrelevant – could be used to facilitate expression judgements.

By the structural-reference hypothesis it is supposed that different routes to face processing are interdependent rather than subserved by functionally and neuroanatomically distinct systems, and that dimensions which are more structurally integral to the face act as referents for the processing of those that are less so – or, that more invariant dimensions act as a referent for more variant dimensions. This logic is akin to that of Garner himself (Garner, 1974), who reasoned that asymmetric patterns of interference were logical in nature - such that X can imply the existence of Y where Y does not imply the existence of X.

In terms of identity and emotional expression, an emotional expression implies an individual (with an identity) expressing it, whereas an individual does not imply an emotional expression as strongly. Extending this logic, Ganel et al. (2004) used the structural-reference hypothesis to explain their findings of bidirectional interference between emotion and identity when faces were familiar, since the structural representation of a familiar face should be richer and more detailed than for an unfamiliar face.

This allows the observer to be more sensitive to structural changes such as those that occur in the changing of an emotional expression. To relate this back to Garner's idea, sufficient knowledge and experience of a person's identity allows an emotional expression which is commonly expressed by that person to stand as a reference for the identity itself – that is, one can process identity information using the emotional expression as a structural referent, since the emotional expression displayed is idiosyncratic to the individual. Y can imply the existence of X if we are familiar enough with both terms.

The structural-reference hypothesis proposes that more structurally integral and invariant dimensions can act as referents for those that are less so – that perceivers can use the underlying structure of a face as a reference for processing information that varies as a function of that structure. Structural information – which arises from the "actual physical shape of the face" (Bruce et al., 1993) – is an important aspect of configural processing, argued to describe the connection between parts or features (Reed, Stone, Grubb & McGoldrick, 2006).

The structural-reference hypothesis only seems to have been applied to studies comparing facial identity and emotional expression (e.g. Soto & Wasserman, 2011; Szweigman, 2020; Wang, Fu, Johnston & Zan, 2013), but may be applicable to other pairs of dimensions. Here, I submit the dimensions of sex and age to the Garner speeded classification task – a pairing of dimensions which has received comparatively little attention in the literature concerned with integrality. As discussed above, this may be because their relationship is not easily predicted in terms of existing face processing models.

An arguably under-investigated distinction between sex and age processing pertains to the relative reliance upon configural and featural information in making categorisation judgements about them. As discussed above (Section 1.6), sex categorisation may depend more on configural processing than age, and age categorisation may depend more on featural processing than sex. Sex processing may also require more attentional resources than age processing (see Wiese et al., 2008, 2012). This latter fact can be inferred from the common finding that sex judgements typically take longer than age judgements in selective attention experiments (Fitousi, 2020; Quinn & Macrae, 2005; Wiese et al., 2008, 2012). Below, I discuss the difference between configural and featural processing in terms of the concept of templates.

4.2.3 Templates

It has been suggested that configural processing may rely on template-like structures (Bartlett & Searcy, 1993; Farah et al., 1998; Tanaka & Farah, 1993; Yuille, 1991). If this is the case, one might argue that representations of sex rely on such *templates*. The “somewhat embattled” term *template* (Tanaka & Farah, 1993) is common in thinking about face perception, it refers to a representational structure relating to a specific object - or level of analysis of an object - which develops with visual experience over time. Templates are typically viewed as operating upon perception in a top-down manner – as pre-existing higher-order representations of a given object which are composed of lower-order representations.

The use of templates in processing means that the specific features making up a template do not need to be explicitly represented (Tanaka & Farah, 1998). In this sense, they are heuristics which negate the need for all aspects of the object to be processed – like Face Recognition Units in Bruce and Young’s model of face processing (Bruce & Young, 1986). Yuille (1991) proposed that there exist global and featural templates – ranging from the level of the whole face to individual features. In agreement with this, Tanaka and Farah (1998) conceive of a hierarchy of templates, with each level of the hierarchy being decomposable into its constituent elements.

Tanaka and Farah’s (1993) hierarchy ranges on the axis of “holistic” to “featural” processing in faces (with “holistic” being deemed a special case of “configural” by some researchers (e.g. Reed, Stone, Grubb & McGoldrick, 2006; Piepers & Robbins, 2012)). They argue that this dichotomy need not be strict, with processing of both kinds likely being utilised to differing extents with relation to different types of objects and under different circumstances. Their work suggests that face *recognition* relies more on holistic (or configural) processing strategies than those used to recognise other types of objects, but it still remains unclear whether and to what extent the classification of different dimensions of the face differentially rely on such processes, and how potential differential processing strategies interact in face perception.

Yuille (1991) argued that the bottom-up or featural strategy of face processing starts at the most local level and works towards the global, and the top-down or configural strategy starts at the most global level and works towards the local. Investigating these strategies, Brunelli and Poggio (1993) found that template-based (top-down) processing was more accurate and simpler to instantiate in computer-based face recognition, but that feature-based (bottom-up) processing could allow for faster recognition due to smaller memory requirements – since parts can be represented using less information than wholes. This is consistent with the general idea that bottom-up processing is faster than top-down processing (e.g. Theeuwes, 2018).

If age is more reliant on featural processing from the bottom-up, and sex with configural processing from the top-down, this may explain why age processing is typically found to be faster than sex processing (Fitousi, 2020; Quinn & Macrae, 2005; Wiese et al., 2008, 2012). This also explains the findings of Schweinberger et al. (1999) and Ganel et al. (2004) that identity - which may rely on more configural processing (Tanaka & Farah, 1998) - is typically slower than emotion processing, which is relatively more likely to rely on featural processing (Beaudry et al. 2012; Bombari et al. 2013)

Rather than sex processing relying solely on configural information and age processing solely on featural information, we might say that representations of sex and age occupy different positions on the *configural processing continuum* (Reed et al., 2006), with sex being closer to the configural end, and age to the featural end. Rather than relying upon templates at one specific level of the hierarchy, they may typically rely on a range of templates skewed to one end or the other. This would mean that sex perception is more reliant on *configural templates*, and age more on *featural templates*.

Marr argued that particular features of an object can be thought of as embedded within the overall structure of an object (Marr, 1982). If we take the structure of the face to be a configural property, we can say that features are embedded within the configuration of the face – or, in terms of a representational hierarchy, that featural templates are embedded within configural templates. Applying this logic to the current question of the perceptual interaction of sex and age in face processing, we might expect that irrelevant changes in sex affect age judgements more so than vice versa.

This is because changes to age information have more of an effect on featural templates which are embedded within configural templates, having less of an effect on the underlying configural templates in which they are embedded. Irrelevant changes in sex, however, change the underlying configural templates within which are embedded the featural templates relied upon in age processing. Irrelevant variation in sex information may affect age processing more so than vice versa, as changes at the configural level may disrupt featural processing more so than vice versa.

Such thinking is consistent with the structural-reference hypothesis (Ganel et al. 2004), according to which sex should act as a structural referent for age processing. Representation of sex may be more stable (Atkinson et al. 2005), “basic” (Schweinberger et al., 1999) and structurally integral to face processing than age representation. We can argue for these points intuitively, since males and females age in different ways, meaning that sex should be able to be used as a reference for age in a way that age cannot be used as a reference for sex. Age also changes irreversibly and constantly over the lifespan, but sex doesn’t, and age changes idiosyncratically in a way that sex doesn’t – making representations of age more unstable. Taking these ideas together, we should expect that age varies as a function of sex rather than vice versa – as emotional expression varies as a function of identity but not vice versa – and thus that irrelevant variation in sex should affect age discrimination more so than vice versa.

4.2.4 Speeded classification of sex and age

A recent example of research which has compared sex and age in a speeded classification task was carried out by Fitousi (2020), who investigated the perceptual integrality of race, sex and age – three relatively invariant dimensions. Fitousi systematically tested the three possible dyadic combinations of these dimensions, first comparing sex and race, then age and race, then sex and age. Fitousi found no Garner effects between any of these pairs of dimensions but did find both *redundancy* and *Stroop* effects – with contemporary speeded-classification tasks typically being designed to permit the measurement and analysis of such effects.

Redundancy and Stroop effects are measurable in selective attention experiments using the speeded classification task because, in such experiments, dimensions typically vary between two values (“old” and “young” in the case of the dimension of age, and “female” and “male” in the case of sex, for example). Redundancy can essentially be analogised with correlation - where there are correlational aspects of perceptual structure between given values across two dimensions, we can say that there is some redundancy between the structure of these values (Maddox, 1992). In such conditions, the processing of information relating to one value leads to facilitated processing of the other value - since some of the information is correlated, and some of the perceptual work has already been done in processing the first value.

Where such facilitation is found relative to the baseline, we refer to it as a *redundancy benefit.* In speeded-classification tasks concerned with face perception, redundancy benefits are typically investigated by including a “positively correlated” block in which all of the stimuli presented exhibit values thought to be correlated across dimensions. These blocks are complemented by a “negatively correlated” block, in which the stimuli exhibit the opposite values to those thought to be correlated - the expectation being that the negative correlation between these values will result in an inhibition of performance relative to the baseline. Such inhibition is known as a *redundancy cost.*

The other effect often investigated is a Stroop effect. The Stroop effect (Stroop, 1935) is supposed to occur when some semantic relationship exists between the *values* of different dimensions such that they interact with one another in processing. With this, Stroop effects are similar to redundancy effects, and they can be measured by comparing performance between the positively and negatively correlated blocks or by analysing performance in the orthogonal block between stimuli which are and are not correlated. Given that the redundancy and Stroop analyses are concerned with the interaction of dimensional values – unlike the Garner analysis which is concerned with dimensions themselves – the redundancy and Stroop effects can be described as “value-level” effects whilst the Garner effect can be described as a “dimension-level” effect.

Any attempt to find a Stroop effect ought to be carried out with regards to “Stroop stimuli” (Algom, Chajut & Lev, 2004), those for which there can be said to be some compatibility or a logical relationship between values across dimensions. The classic example of this might be words for different colours (e.g. “red” and “green”) written in different colours (e.g. red and green); there is a logical relationship between the semantic meaning of the word and the colour in which it is written. Such logical relationships have been noted for different combinations of social dimensions of the face. Kloth et al. (2015), found that the effects of increasing age differentially affected sex judgements about male and female faces - such that female judgements were faster in young faces compared to male judgements, and male judgements were faster in old faces compared to female judgements.

Johnson et al. (2012) found dependencies between sex and race using artificial faces to morph from Black, through Caucasian, to Asian faces. They found that female judgements were facilitated across this trajectory and male judgements were inhibited, suggesting a perceptual relationship between Black and male faces, and Asian and female faces. According to such findings, young female, old male, Asian female and Black male faces can be defined as “Stroop stimuli” - those for which we would expect to find redundancy benefits based on a logical relationship between the values of the given dimensions.

Findings such as those of Kloth et al. (2015) and Johnson et al. (2012) can be thought of in terms of the shared signal hypothesis (e.g. Adams & Kleck, 2003), by which it is supposed that redundant values share some underlying value – such as approach-avoidance in the case of redundancy between angry expressions and direct gaze on the one hand, and fearful expressions and averted gaze on the other. It is not presently clear what shared signal underlies young female and old male, or Black male and Caucasian female faces (as shown by Kloth et al. (2015) and Johnson et al. (2012) respectively). All that seems evident is that these faces are processed more efficiently than faces exhibiting other combinations of values across these dimensions.

In terms of redundancy and Stroop effects, Fitousi’s (2020) found there were redundancy effects and Stroop effects in the sex and race task. Performance was facilitated when targets were Black males or Caucasian females compared to when they were Black females or Caucasian males - replicating the findings of Johnson et al. (2012). These Stroop and redundancy effects were bidirectional, affecting both sex and race judgements equally. Fitousi also conducted an experiment using Caucasian and Asian faces instead of Black and Asian faces and found no such effects. Looking at race and age, he found no redundancy or Stroop like effects comparing Old and Young Black and Caucasian faces.

Looking at age and sex, Fitousi (2020) found redundancy benefits and costs for the sex task only, suggesting that age related information could interfere with sex judgements but not vice versa. This finding is consistent with that of Quinn and Macrae (2005) who also investigated the relationship between sex and age. They found in a repetition priming experiment that participants were faster to make a subsequent sex judgement on a face which they had previously made an age judgement about, but were not faster to make an age judgement if they had previously made a sex judgement. This suggests that age information must be accounted for in sex processing, but not vice versa. In the orthogonal blocks, Fitousi found bidirectional Stroop interference - replicating Kloth et al. (2015) - finding facilitated performance for young females and old males relative to old females and young males.

Like Fitousi (2020), Quinn and Macrae's (2005) finding suggested that participants were sensitive to age related information about a stimulus when making a sex judgement but not an age judgement - that sex judgements were more susceptible to interference from age information than vice versa. Following on from this experiment with a speeded-classification task, Quinn and Macrae supported the findings of their repetition priming task by showing asymmetrical Garner interference, such that sex judgements were inhibited by irrelevant variation in age but that age judgements were not inhibited by irrelevant interference in sex.

Fitousi (2020) pointed out issues with the Quinn and Macrae's speeded-classification task, highlighting that they did not include correlated blocks nor investigate Stroop effects. Fitousi also argues that age was "far more discriminable than sex", rendering their conclusion of asymmetrical integrality "dubious". This latter criticism is presumably based on the fact that age discrimination was shown by Quinn and Macrae (2005) to be significantly faster than sex discrimination, and that they did not correct for this - although Fitousi (2020) went on to find the same differences in latency and also did not correct for it. As has been shown in the cases of Atkinson et al. (2005) and Schweinberger et al. (1999), these differences in task-difficulty cannot be said to wholly account for asymmetric integrality between the dimensions.

An important aspect of Fitousi's (2020) findings were that he found Stroop effects but no Garner effects when looking at the dimensions of age and sex and sex and race. His rationale for finding this pattern of results relies on the supposition that Stroop effects occur prior to Garner effects, a finding reported by Boenke et al. (2009) in which they found a difference in time course between Stroop and Garner effects using ERP. They argued that Stroop effects are sensitive to interference within presentations and that Garner effects are sensitive to interference between presentations.

Boenke et al. (2009) suggest that the Garner effect involves expectancy and memory – and thus involves more of a semantic component, occurring later in the perceptual processing stream. They argued on the basis of their ERP results that the Stroop effect has both a visual and a cognitive or "control" component – manifesting in occipital and fronto-central areas. The Garner effect, on the other hand, only manifested in the fronto-central areas – suggesting that there is no visual component to the effect. Stroop effects in the absence of Garner effects have been reported only rarely (Fitousi, 2016, 2020; Patching & Quinlan, 2002; Van Leeuwen & Bakker, 1995), which Boenke et al. (2009) attributes to the relatively large timescale in which the temporal signatures of these effects are thought to overlap.

In keeping with this explanation from Boenke et al. (2009), Fitousi (2020) championed a holistic-to-analytic view of perceptual processing by which stimuli are initially processed as a whole - accounting for the presence of Stroop effects - and are only later processed in terms of their parts - resulting in the absence of Garner effects. At the early, holistic stage, the values along each dimension interact, but by the time the dimensions are processed analytically the dimensions are no longer in interaction - according to Fitousi's view.

As far as I'm aware, Quinn and Macrae (2005) and Fitousi (2020) are the only researchers to have conducted speeded-classification tasks comparing the dimensions of sex and age - and only Fitousi carried out redundancy and Stroop analyses. These researchers' findings - though partially consistent with one another – are inconsistent with the predictions of the popular Haxby et al. (2000, 2002) model, according to which sex and age should be bidirectionally integral, or their asymmetrical integrality should be in the other direction such that sex interferes more with age than vice versa – since sex is arguably less variant than age.

Generally speaking, investigations of perceptual interference are lacking for this pair of dimensions, and what little research has been done is either incomplete or is inconsistent. Seeking to address this issue, the current experiment was designed to test the perceptual integrality of sex and age using the Garner speeded-classification task - using the full contemporary design in line with the work of Fitousi (2020). Unlike Fitousi, in this study experimental task (sex or age) is a between-subjects, rather than a within-subjects, factor.

This was in order to mitigate any possible task-switching effects on the basis that such effects might be asymmetrical – particularly if there is an asymmetric pattern of Garner interference between these dimensions. It was also to avoid the possibility of effects emerging as a function of response level competition. That is, if a participant is not aware of the response mapping for the irrelevant dimension, this response mapping cannot interfere with processing. This means that, if Garner, Stroop or redundancy-based interference is found, we can assume that the interference occurs at the perceptual level rather than the response level.

I used three stimulus sets instead of one in the experiment. A criticism of Schweinberger et al. (1999) levelled by Ganel et al., (2004) was that they used different stimuli for their different tasks – allowing for the possibility that the difference in effects between tasks had to do with differences in the stimulus properties. Fitousi (2020), similarly, used two different stimulus sets across dimensions – mixing them within experiments. Quinn and Macrae (2005) did not specify which stimulus set they used, reporting only that they presented “graphics files depicting greyscale images of faces”. Reasoning that inconsistencies in stimuli within and across experiments may contribute to inconsistencies in results, the current experiment uses three different sets of stimuli between participants (see chapter 3 for more details about the stimuli used).

Exploring the relationship between task-difficulty and each type of interference (Garner, Stroop and redundancy) across the three different stimulus sets should help us to understand whether and to what extent any differences arise as a function of dimensional discriminability. For instance, finding differences in task-difficulty between sets as well as corresponding differences in interference effects would suggest a role of discriminability, whilst differences in task-difficulty paired with consistencies of interference effects would suggest that the interference effects emerged as a function of the organisation of the representational structure of the dimensions and values under investigation – independently of the stimulus properties.

4.3 Methods

Please refer to the *General Methods* section (Chapter 3) for more general information about this experiment.

4.3.1 Participants

A total of 321 participants took part in this experiment. Demographic information was recorded for 254 participants. Of those, 92 (36.22%) reported their sex as male, 109 (42.91%) reported their sex as female, and 53 (20.87%) did not report their sex. The mean age of the participants who provided age information (254) was 25.99 (min = 18, max = 56). A total of 19 participants were excluded as outliers in terms of overall accuracy or RT, or for too many invalid responses (pressing keys other than the designated response keys) (5.92%), leaving 302 participants. Following these exclusions, 0.49% of individual trials were removed as invalid responses, and responses that were < 200ms or > 2500ms were removed (0.91%).

Unfortunately, I was unable to match the recorded demographic data to specific participants, and so I don't know the demographics of participants who were included in the study. Participants were paid £2 (Median duration = 10:55). All procedures were granted ethical approval by the Research Ethics Committee of Bangor University's School of Psychology. The experiment was created using Jatos (https://www.jatos.org/) and p5 (https://p5js.org/).

N	SEX	AGE	TOTAL
FACEGEN	49	49	98
MPI	52	54	106
LIFESPAN	47	51	98
TOTAL	148	154	302

Table 4.1: Number of participants per *task* and *set*

4.3.2 Design and Procedure

In a mixed design there were two between-subjects variables, *task* (judge faces on either Age or Sex) and *set* (Facegen, MPI or Lifespan), and one within-subjects variable, *condition* (Baseline, Orthogonal, Pos Corr, Neg Corr) (see Figure 4.2). Within *condition*, there were two Baseline blocks for each *task*, one for which the irrelevant-dimension was held constant at one value and one for which it was held at the other. In the Orthogonal block, values on the irrelevant dimension and the relevant dimension varied such that each combination of values was represented an equal number of times. In the positively correlated blocks (Pos Corr), the values of the relevant and the irrelevant dimensions were correlated according to “compatibilities” noted in the literature (Fitousi, 2020; Kloth et al. 2015), specifically, all stimuli were either Male and Old or Female and Young. In the negatively correlated blocks (Neg Corr), the values of these dimensions were the opposite of those in the Pos Corr blocks such that they were all Female and Old or Male and Young .

Immediately before the experimental phase began, I estimated the viewing distance of participants using a Virtual Chinrest approach (Li, Joo, Yeatman, Reinecke, 2020). This procedure involves two steps. First, participants matched the size of an on-screen credit card with a real card of equivalent size. This procedure, along with screen pixel measurements, provides an estimate of the density of the participants’ displays in pixels/mm. Next, participants took part in a blind-spot task, which involves covering one eye and focusing on a black square, and then pressing the space bar when a horizontally-translating red circle moved out of visibility. Engaging in these tasks allowed estimation of the approximate viewing distance of the participant, based on an average angle of eccentricity of the blind spot of about 13˚. On this basis, the software for the present experiment was coded to present faces at a varying pixel size that equated to roughly 5° visual angle across participants.

All participants then took part in each of the five blocks, drawn from the four *conditions*. In keeping with previous research in the field (Fitousi, 2020) blocks were presented in a random order. Each block contained 36 trials. Two Orthogonal blocks were presented to compensate for there being two Baseline blocks and two Correlated blocks. Trials were presented in a random order within blocks.

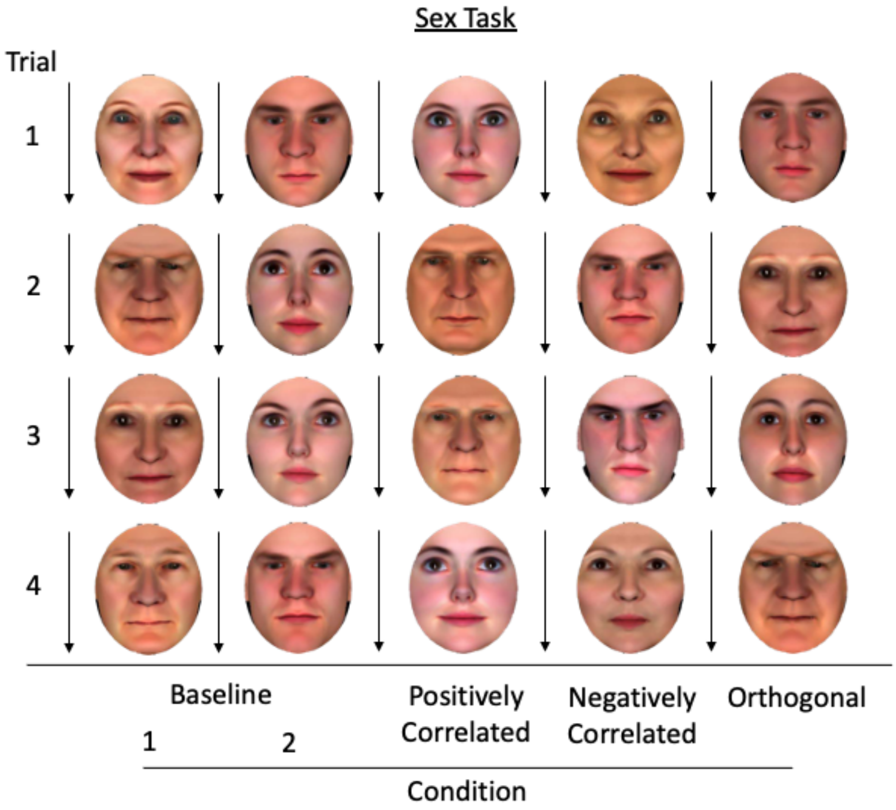


Figure 4.2: Example of four trials of the Sex *task* using the Facegen *set* for each of the five blocks. In each of the Baseline *conditions,* the relevant dimension (Sex) varied over trials, but the irrelevant dimension (Age) was held constant at one or the other level; in this example , Age is held constant at Old for the first, and Young for the second Baseline block. In the Positively Correlated block, the level of the irrelevant dimension is compatible with the given level of the

relevant dimension (Young Female and Old Male), whereas in the Negatively Correlated block, the irrelevant and relevant dimensions are incompatible (Young Male and Old Female; see Kloth et al., 2015). In the Orthogonal block, the levels of the relevant and irrelevant dimensions were counterbalanced such that each of the four combinations of those levels appeared equally often.

Participants were informed that each trial would consist of the presentation of a face, and that their task was to respond as quickly and as accurately as possible to either the Sex or Age of the face (depending on the *task*). The fixation cross and the target face were always presented centrally and in the same location. The fixation cross was presented for 300ms and then the faces were presented for a maximum of 3000ms, or until a response was registered. There was a brief inter-trial pause of 84ms.

4.4 Results

4.4.1 Control Analyses

Control analyses were carried out in order to inform the interpretation of the main experimental analyses. The first of these analyses examined differences in IES (Inverse Efficiency Scores computed through taking the mean RT on accurate trials, see *General Methods*) between the Baseline *conditions* for each *task* and *set*. I investigated differences between the Baseline *conditions* by collapsing across *tasks* and *sets* and using a paired samples t-test, which revealed no significant differences ($t\ (301) = 0.83$, $p = 0.41$, $d = 0.03$). This means that performance in the Baseline blocks did not vary as a function of which task-irrelevant value was kept constant.

A second control analysis investigated differences between the *tasks* and *sets* collapsed across all five *conditions*. This was achieved using a two-way between-subjects ANOVA with *task* and *set* as factors. This analysis revealed that the Age *task* was performed more efficiently than the Sex *task* ($F(1,296) = 28.68$, $p < 0.001$, $h^2 = 0.09$; Sex M = 876.89 (12.48), Age M = 801.34 (9.05)). This was qualified by an interaction between *task* and *set* ($F(2,296) = 21.64$, $p < 0.001$, $h^2 = 0.13$). Follow-up analyses revealed a significant difference between *tasks* for each *set.* For the Facegen *set,* performance was significantly better in the Sex compared to the Age *task* ($t(97) = 2.33$, $p = 0.02$, $d = 0.47$; Sex M = 796.09 (18.43), Age M = 853.45 (16.24)). In the other two stimulus sets, performance was significantly better in the Age compared to the Sex *task:* MPI (t(92) = 5.94, p < 0.001, d = 1.16; Sex M = 906.61 (19.64), Age M = 763.76 (13.89)); Lifespan: (t(81) = 5.49, p < 0.001, d = 1.13; Sex M = 935.69 (21.92), Age M = 791.06 (14.57))

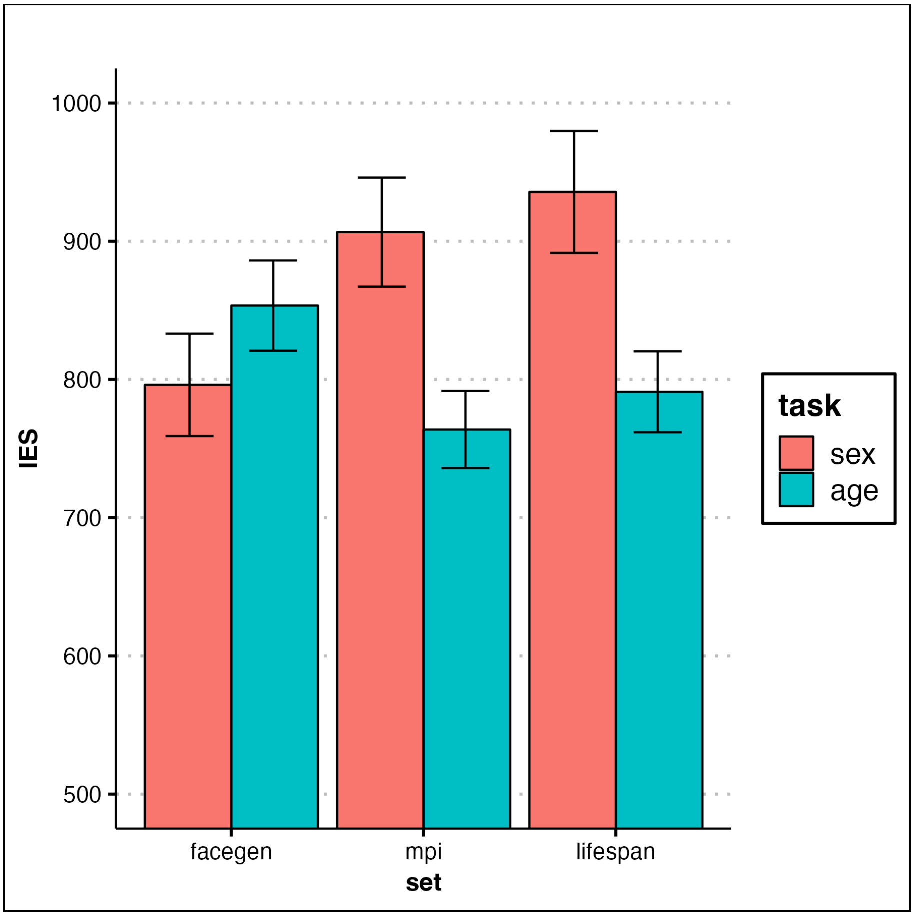


Figure 4.3: Inverse efficiency scores for each *set* and *task* collapsed across *conditions.* IES is calculated as the mean reaction time on accurate trials divided by the proportion of correct trials. Error bars represent 95% confidence intervals and the asterisks indicate significant differences at a threshold of p <0.05 or lower. These conventions hold for all subsequent figures.

4.4.2 Garner Interference

Garner interference was measured as a difference score calculated by subtracting performance in the Baseline *condition* from performance in the Orthogonal *condition.* A positive difference score would tell us that performance was worse in the Orthogonal *condition* relative to the Baseline, indicating Garner interference. Participants whose difference score (averaged across the two *conditions*, Baseline and Orthogonal) fell outwith 2.5 SDs of the mean difference score (collapsed across *task* and *set*) were excluded from the Garner interference analysis (n = 10).

To assess Garner interference and its possible modulation by *task* and by stimulus *set*, difference scores were submitted to a two-way between-subjects factorial ANOVA with *task* (Age or Sex), and *set* (FaceGen, MPI or Lifespan) as factors. This analysis showed significant Garner interference overall (M = 28.02 (3.89), interpreted as a y-intercept significantly different from zero (F(1,286) = 44.65, $p < 0.001$, $h^2 = 0.135$), meaning that performance was significantly worse in the Orthogonal compared to the Baseline *conditions*. There was also a significant main effect of *task* (F(1,286) = 5.28, p = 0.02, $h^2 = 0.018$), such that Garner interference was significantly greater for the Age *task* (M = 37.26 (5.39)) compared to the Sex *task* (M = 18.54 (6.48)). Follow-up t-tests revealed that Garner interference was significant for both the Sex *task* (t(143) = 2.86, p = 0.005, d = 0.24) and the Age *task* (t(147) = 6.91, $p < 0.001$, d = 0.57). This tells us that Garner interference occurred for both *tasks* but was asymmetrical, being larger for the Age than for the Sex task.

There was no significant main effect of *set* (F(2,286) = 2.36, p = 0.09, $h^2 = 0.016$) and no interaction between *task* and *set* (F(2,286) = 0.19, p = 0.82, $h^2 = 0.001$), from which I infer that the pattern of asymmetrical Garner interference held in terms of direction and magnitude across *sets.*

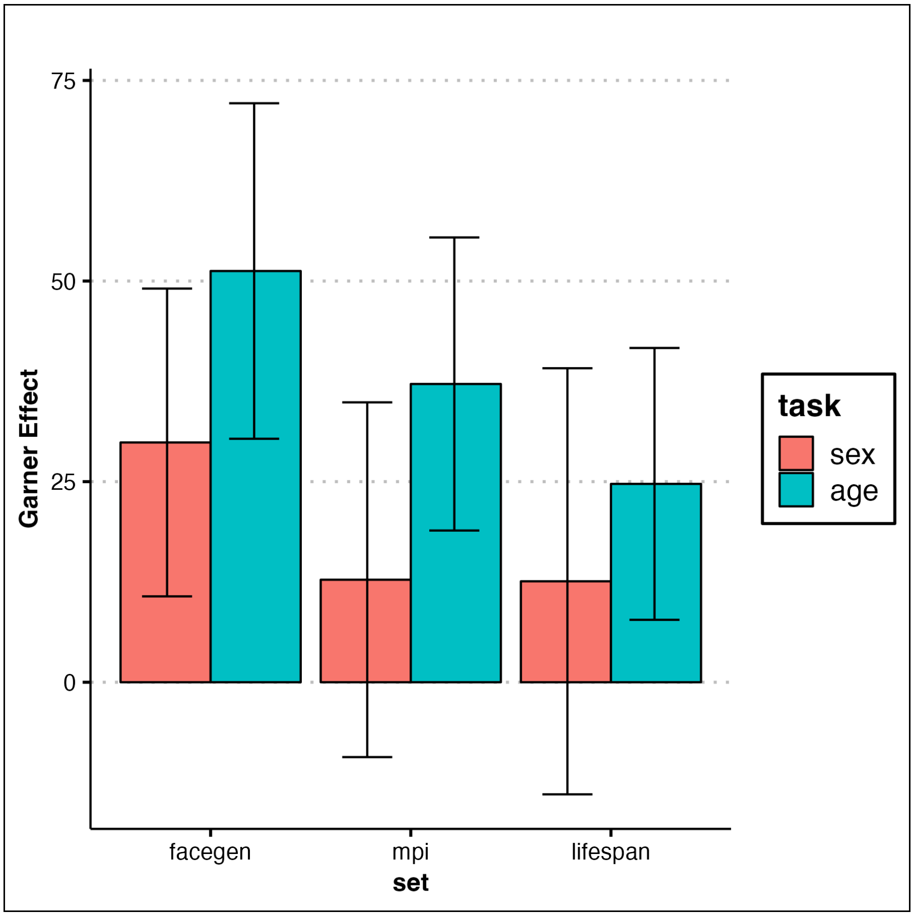


Figure 4.4: Garner interference as a function of *task* and *set* measured as the difference in inverse efficiency scores (IES) between the Baseline and Orthogonal conditions.

4.4.3 Stroop Effects

Difference scores were calculated by subtracting the Compatible from the Incompatible configurations (see chapter 3) within the Orthogonal block. Stroop effects and their possible modulation by *task* and by stimulus *set* were assessed by submitting difference scores to a 2x3 two-way between-subjects ANOVA with *task* and *set* as factors. A positive difference score indicates better performance in the Compatible compared to the Incompatible condition, and is interpreted as the presence of Stroop effects. Participants whose absolute difference score fell outwith 2.5 SDs of the absolute mean difference score (collapsed across *task* and *set*) were excluded from the Stroop effects analysis (n = 12).

This analysis showed significant, positive, overall Stroop effects (M = 71.13 (5.89), interpreted as a y-intercept significantly different from zero (F(1, 284) = 137, $p < 0.001$ $h^2 = 0.325$). This means that stimuli with Compatible values across dimensions were categorised more efficiently than those with incompatible combinations of values. This effect was significantly modulated by an interaction between *set* and *task* (F(2 284) = 12.45, $p < 0.001$, $h^2 = 0.08$).

Follow-up analyses showed a significant difference between *tasks* for both the Facegen (t(83) = 4.52, $p < 0.001$, d = 0.95; Sex M = 40.65 (13.68), Age = 137.52 (16.52)) and the MPI *sets* (t(77) = 2.7, p = 0.009, d = 0.53; Sex M = 93.79 (17.5), Age M = 40.05 (9.55)), but not for the Lifespan *set* (t(63) = 0.48, p = 0.63, d = 0.1; Sex = 55.38 (21.87), Age = 67.09 (10.59)). It should be noted that the interaction is such that the Stroop effect was greater for the Age *task* than the Sex *task* for the Facegen *set* and vice versa for the MPI *set*.

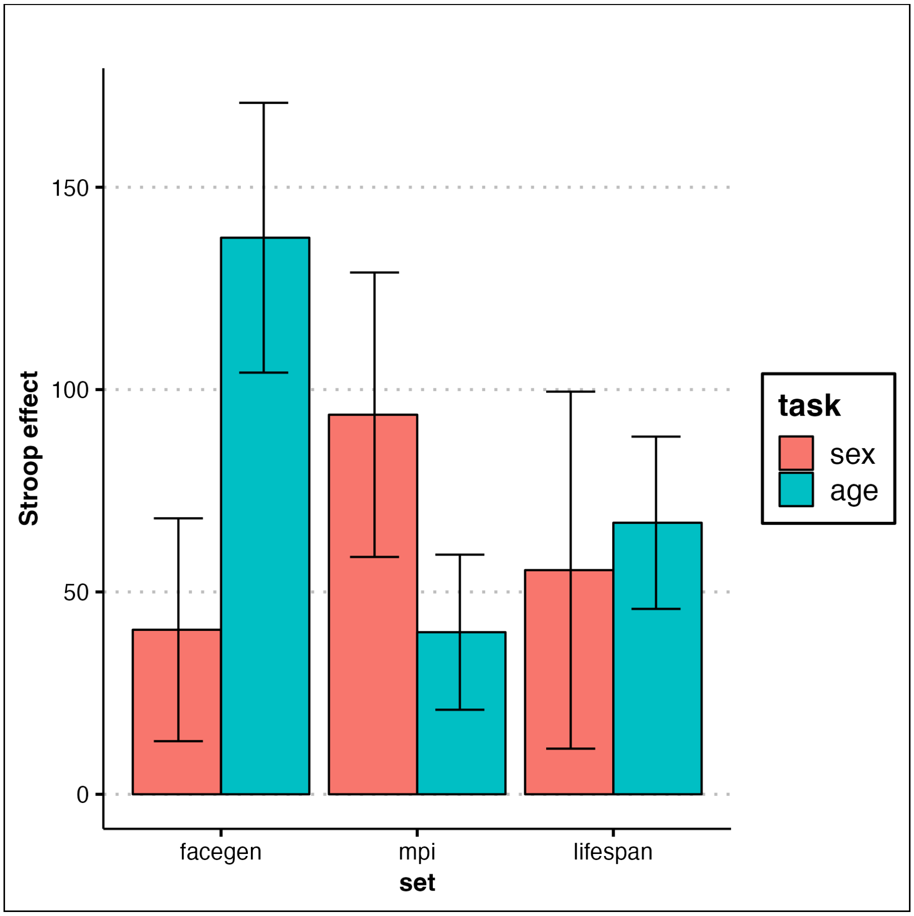


Figure 4.5: Stroop effects as a function of *task* and *set*. Positive values indicate Stroop effects: increases in performance owed to compatibility of values across stimulus dimensions (e.g. Old

and Male, or Young and Female) relative to the incompatible trials (Old and Female, or Young and Male).

4.4.4 Redundancy Benefits

To determine the presence and extent of redundancy benefits – a facilitation of performance associated with the positive correlation between Sex and Age (Male/Old and Female/Young) – a difference score was calculated by subtracting the Positively Correlated *condition* from the Baseline *condition.* A positive difference score would be interpreted as redundancy benefits emerging from the compatibility between levels of the experimental dimensions compared to each *task's* respective control condition. Participants whose difference score fell outwith 2.5 SDs of the absolute mean difference score (collapsed across *task* and *set*) were excluded from the redundancy benefits analysis ($n = 10$).

This analysis showed a significant Redundancy benefit ($M = 42.47$ (4.65)), interpreted as a y-intercept significantly different from zero ($F(1,286) = 78.71$, $p < 0.001$, $h^2 = 0.216$). There was a significant main effect of *task* such that there were higher redundancy benefits for the Sex compared to the Age *task* ($F(1,286) = 6.86$, $p = 0.01$, $h^2 = 0.023$, Sex $M = 56.57$ (7.63), Age $M = 29.66$ (6.57)) but no main effect of *set* ($F(2,286) = 0.52$, $p = 0.59$, $h^2 = 0.003$). There was a significant interaction between *set* and *task* ($F(2,286) = 7.41$, $p < 0.001$ $h^2 = 0.05$).

Follow-up analyses showed a significant difference in the redundancy benefits between *tasks* for the MPI *set* only ($t(88) = 4.57$, $p = < 0.001$, $d = 0.91$; Sex $M = 73.26$ (1.67), Age $M = 7.6$ (1.13) with a marginal effect being observed for the Lifespan *set* ($t(72) = 1.92$, $p = 0.06$, $d = 0.41$; Sex = 57.81 (2.45), Age = 21.55 (1.44). For both *sets,* greater redundancy benefits were observed in the Sex relative to the Age *task*. It should be noted that the non-significant difference between tasks for the Facegen *set* ($p = 0.19$) was in the opposite direction.

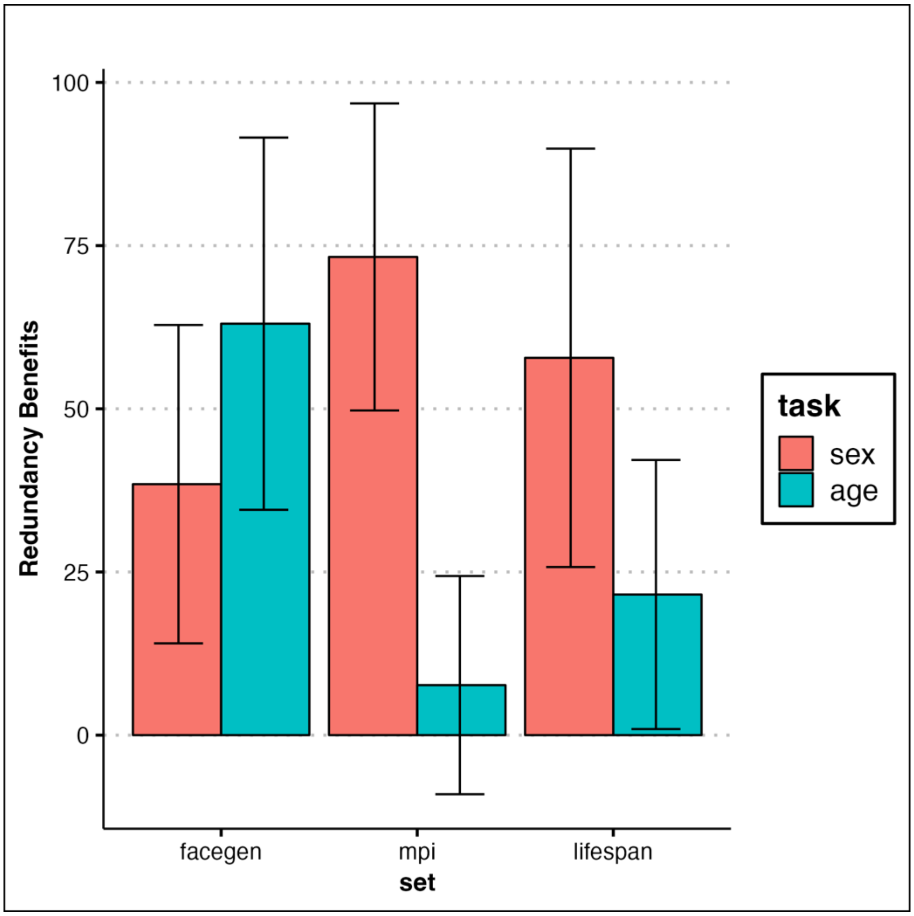


Figure 4.6: Redundancy benefits as a function of *task* and *set*. Positive values indicate redundancy benefits, which are improvements in performance owed to compatibility of values across stimulus dimensions (e.g. Old and Male, or Young and Female).

4.4.5 Redundancy Costs

To determine the presence and extent of redundancy costs – a disruption of performance associated with the negative correlation between Sex and Age (Male/Young and Female/Old) – a difference score was calculated by subtracting the Baseline *conditions* from the Negatively Correlated *condition.* A positive difference score would be interpreted as redundancy costs emerging from the incompatible values between levels of the experimental dimensions compared to each *task's* respective controls. Participants whose absolute difference score fell outwith 2.5 SDs of the absolute mean difference score (collapsed across *task* and *set*) were excluded from the redundancy costs analysis (n = 10).

This analysis did not show a significant overall redundancy cost (M = 8.66 (0.29)), as interpreted by a y-intercept not significantly different from zero (F(1,286) = 2.1, p = 0.14, h² = 0.007). However there was a significant main effect of *task* such that there were bigger redundancy costs for the Age compared to the Sex *task* (F(2,286) = 5.55, p = 0.02, h² = 0.019; Sex M = -5.3 (7.63); Age M = 22.23 (6.57)). Follow-up t-tests revealed that there was a significant redundancy cost for the Age *task* (t(147) = 3.15, p = 0.002, d = 0.26) but not for the Sex *task* (t(143) = -0.59, p = 0.56, d = -0.05). There was also a significant interaction between *task* and *set* (F(2,286) = 4.82, p = 0.009, h² = 0.033).

Follow-up analyses revealed that there was a significant difference between *tasks* for the MPI (t(95) = 3.07, p = 0.003, d = 0.6; Sex M = -27.03 (2.17), Age M = 33.09 (1.6)) and the Lifespan *set* (t(65) = 2.14, p = 0.04, d = 0.46; Sex M = -21.1 (2.61), Age M = 20.71 (1.28)) such that costs for Age were significantly larger than those for Sex. In fact, for these *sets*, the redundancy cost was negative, although not significantly different from zero (both > 0.09), indicative of redundancy benefits. For the Facegen *set*, the difference between *tasks* followed the opposite trend such that costs were higher for the Sex compared to the Age *task* – although this difference was not significant (t(84) = -1.09, p = 0.28, d = -0.23; Sex = 32.70 (1.82); Age = 10.64 (2.39)).

It is possible that the failure to find a significant overall redundancy cost occurred because some of the values were negative – as discovered through the interaction between *task* and *set*. In order to understand the result more deeply, I therefore carried out one-sample t-tests for each level of *task* for each *set*. These revealed that there were significant positive differences from zero (indicating redundancy costs) for the Facegen *set* in the Sex *task* (t(47) = 2.59, p = 0.01, d = 0.37) but not the Age *task* (p = 0.5), for the MPI *set* in the Age *task* (t(53) = 2.8, p = 0.006, d = 0.38) and in the Lifespan *set* for the Age *task* (t (49) = 2.29, p = 0.03, d = 0.32).

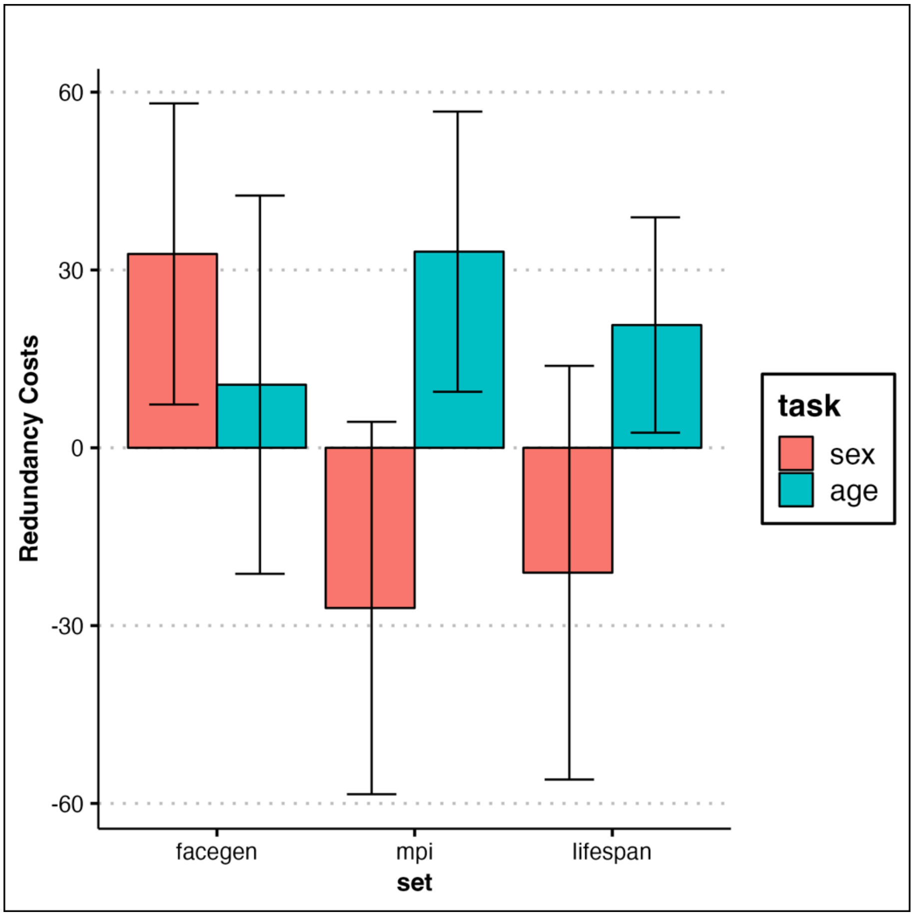


Figure 4.7: Redundancy costs as a function of *task* and *set*. Positive values indicate redundancy costs, which are defined as decreases in performance owed to incompatibility of values across stimulus dimensions.

4.5 Discussion

The major finding from the current results is evidence for bidirectional Garner interference between the dimensions of Sex and Age. This interference is asymmetric, such that irrelevant Sex information interferes with Age processing to a greater degree than vice versa. The Garner effect was consistent across stimulus *sets* despite differences between these *sets* in terms of baseline task-difficulty – suggesting that the Garner effect is not modulated by baseline task-difficulty and therefore cannot be accounted for in terms of the SoPA (see Graham & LaBar, 2007; Melara & Mounts, 1993).

I found evidence for redundancy effects and Stroop effects which show that Old Male and Young Female faces can be processed more efficiently than Young Male and Old Female faces. This interaction between dimensional values appears to be more facilitative than it is inhibitory, given the finding of significant redundancy benefits in the face of weaker or non-significant redundancy costs. The patterns of Stroop and redundancy effects between *tasks* differ between *sets* in a way that is consistent with the differences in task-difficulty between the *sets*, suggesting that these effects are modulated by baseline task-difficulty. This suggests a dissociation between the processes underpinning dimension-level (Garner) and value-level (Stroop and redundancy) effects.

There are a number of similarities and differences between the *sets* – rendering a simple explanation of the results difficult. Fortunately, these similarities and differences are largely consistent, and where I found no significant effects there generally remain suggestive patterns and trends which facilitate interpretation. I will first go over and summarise the key findings from each of the main analyses with reference to the baseline task-difficulty for each *set*. I will then discuss the similarities and differences between the *sets*, before offering an overarching interpretation of the findings.

4.5.1 Garner Findings

The finding of asymmetric Garner interference is particularly interesting considering the relative task difficulty. It has been suggested that asymmetric Garner interference can sometimes be explained by easier tasks interfering with harder tasks (SoPA, Graham & LeBar, 2007; Melara & Mounts, 1993). Looking back at the control analyses (see Figure 4.3) we can see that the Age *task* was easier to perform than the Sex *task* for both photographic *sets*. Despite this, the asymmetric Garner interference occurs in the opposite direction. Therefore, a simple task-difficulty account of asymmetric interference cannot account for the present findings.

The pattern of task-difficulty (Age easier than Sex) is reversed for the Facegen *set*, such that the Sex *task* was easier than the Age *task*. This means that the SoPA of asymmetric interference could apply to the Facegen *set*, however this reasoning is weakened by the fact that the other two *sets* display the same pattern of asymmetry but a different pattern of task-difficulty.

4.5.2 Stroop findings

There was a large, significant, Stroop effect for each of the three *sets*, this was larger for the Age than for the Sex *task* for the Facegen *set*, larger for the Sex than the Age *task* for the MPI *set,* and comparable between tasks for the Lifespan *set*. For the two *sets* with significant differences between *tasks*, we see that the *task* for which there was a greater Stroop effect is the *task* which was more difficult (Age for Facegen and Sex for MPI). Given this consistency, one would have expected to find a larger effect in the Sex than the Age *task* for the Lifespan *set.*

4.5.3 Redundancy effects

I compared performance on Baseline trials with trials in which the faces shown had positively or negatively correlated values across dimensions to investigate redundancy benefits and costs respectively. There were significant redundancy benefits but not costs – although, for the photographic *sets*, there were significant differences between *tasks* in terms of costs. This means that stimuli with compatible values across dimensions reliably and significantly facilitated performance whereas stimuli with incompatible values did not reliably and significantly inhibit performance. We see the same trend in the results emerge for each of the *sets*, greater benefits for the harder task and greater costs for the easier *task.* For both photographic *sets*, the redundancy costs were negative for the harder, Sex *task* – meaning that even when the dimensional values were negatively correlated or incompatible with the given Sex judgement, Age information facilitated performance.

4.5.4 Similarities and differences between Sets

There was significant Garner interference for all three *sets* and this interference was bidirectional – occurring more for the Age *task* than for the Sex *task,* suggesting that irrelevant Sex information interfered with Age judgements to a higher degree than vice versa. There were also significant Stroop and redundancy benefits for all *tasks.* There were no significant redundancy costs for any *set*. The redundancy costs show a similar pattern between the MPI and Lifespan (photographic) *sets*, such that both show redundancy costs for the Age *task* and redundancy benefits for the Sex *task.* That is, even when the values are incompatible, redundant Age information facilitates Sex judgements.

The major differences between *sets* lies in the direction of difference in performance between *tasks* across *sets* for the Stroop and redundancy analyses. These differences are typically split such that the photographic MPI and Lifespan *sets* are similar, but dissimilar to the artificial Facegen *set*. As mentioned above, some of these differences are not significant, but are typically suggestive of a trend and are consistent across analyses within the same *set*.

Differences in baseline task difficulty seem to be represented in the differences between the Stroop and redundancy analyses. That is, there was generally greater Stroop interference and redundancy gains for the harder *task*, and lower redundancy costs for the harder *task.* The main exception to this pattern was the absence of a difference between *tasks* in the Stroop analysis of the Lifespan *set* – wherein the difference bucks the trend with more Stroop interference in the easier *task.* However, this small and non-significant difference may simply be an anomaly rather than a telling data point.

Overall, the differences here noted suggest that the more difficult the *task* is, the greater the extent to which redundancies in the stimulus structure from task-irrelevant but more easily discerned information positively influences decision making.

4.5.5 Garner Interpretation

The result is consistent with the proposal of Boenke et al. (2009) that the Garner effect does not emerge at the visual level, but later in the processing stream. If the Garner effect were modulated visually, we would expect to see differences in its manifestation as a function of task-difficulty. The current results suggest that the asymmetrical Garner effect emerges more so at higher-order semantic levels, rather than at lower-order visual levels of the representational hierarchy. In the current experiment, it cannot be said to be occurring at the response level, since *task* was a between-subjects factor, and thus the participant could not know the response pattern for the irrelevant *task.*

That being said, the emergence of the Garner effect must proceed from the available visual information – information which appears to mitigate both the relative discriminability (difficulty) of the two dimensions, as well as the Stroop effect. The Garner effect could be argued to emerge up-stream from the Stroop effect, and may be informative about the underlying mechanics and functional organisation of the processing systems involved in the perception of sex and age.

I argue that such systems are characterised by a differential reliance on configural and featural processing. In the orthogonal blocks of the Age *task*, there is more irrelevant variation in the configural information within which the featural information – relied upon to make the age judgement – is embedded. In the Sex *task*, one has to process the configural information underlying the featural information in order to make the judgement. Irrelevant variation in the featural information does not interfere as strongly with that judgement. This difference in processing strategy can be conceived in terms of differences in reliance on configural and featural templates for sex and age processing respectively.

According to Garner (1962), the larger the amount of information that needs to be processed, the greater the processing load, leading to less efficient responses. In line with previous findings (Fitousi, 2020; Quinn & Macrae, 2005; Wiese et al., 2008, 2012) I found less efficient responses in the Sex *task* compared to the Age *task,* suggesting that sex discrimination requires a greater processing load than age discrimination. Wiese argued on the basis of the existing evidence that this was because accurate age categorisation could occur in relation to “surface cues”, whilst sex categorisation required a processing of the configuration of facial features. This is in line with Brunnelli and Poggio’s (1993) finding that recognition could occur more quickly using features compared to configural templates, and suggests that sex perception may require more complex and resource-intensive processing than does age perception.

Whilst it has been shown that sex judgements can be made using only specific facial features, such as the nose (Chronicle et al., 1995) the judgement might not emerge from the features of that feature – so to speak - but from the configuration of that feature. This is evidenced by the finding of Chronicle et al. that three-quarter views of the nose permitted sex judgements to be made more accurately compared to frontal views – where less of the 3-D configural information is visually available. That is to say, the configuration and/or the features of a facial feature can be processed. There are parts of parts and wholes of parts, with processing of the former being featural processing and processing of the latter being configural processing.

I found that the pattern of task-difficulty differed between the photographic (MPI and Lifespan) *sets* and the artificial (Facegen) *set* – such that the Age *task* was easier for the former and more difficult for the latter *set* than the Sex *task*. With regards to the Facegen *set*, it may be the case that age is less discriminable, and/or that sex decisions are more discriminable relative to the photographic *sets*. Decreased discriminability of age may have occurred because the stimuli are less familiar, and thus processing based on individual features (which may underpin dichotomic age decisions (George & Hole, 1998; Quinn & Macrae, 2005)) may become a less efficient strategy. Skin texture – a “global feature” involved in age perception (Rhodes, 2009) – is also likely to be less realistic in the artificial compared to the photographic *sets*.

That is, the match between the featural templates (pre-existing representational structures) and the observed features (of the stimulus) is not achieved as efficiently as it would be with photographic stimuli with which the observer has more visual experience. Increased discriminability of sex may occur because of the potentially caricatured differences in the information distinguishing Male and Female Facegen stimuli – information which is presumably configural in nature.

A crucial point in the current Garner results is that Garner interference was uniform between *tasks* across *sets*. That is, despite the difference in task-difficulty for the artificial compared to the photographic *sets*, the pattern of integrality was the same – sex interferes with age more than vice versa. This constitutes further evidence against the SoPA in line with Atkinson et al. (2005) and Schweinberger et al., (1999). The results of the Facegen *set* taken alone would support the SoPA, but the results of the photographic *sets* go against it – since, in those *sets*, the less discriminable dimension interfered with the more discriminable dimension more so than vice versa.

This result calls into question the validity of basing interpretations on data collected using a single stimulus set, and also represents a major challenge to the SoPA. It is not a challenge purely because two *sets* go against it and only one *set* supports it, but because the result is homogenous across *sets* regardless of task-difficulty differences. Finding such a homogenous pattern of interference despite differences in dimensional discriminability may suggest that Garner interference is occurring at the level of perceptual representation rather than at the level of stimulus properties – that is, the interference may arise at more semantic and less visual levels. The effect may be driven from the top-down rather than from the bottom-up. If interference occurred at the level of sensation and was a bottom-up phenomenon, we ought to expect the pattern of integrality to be more dependent on discriminability. Furthermore, the interference cannot be arising at the level of response selection, since *task* was a between-subjects variable.

That being said, it is possible that the observed differences in dimensional discriminability within and between *sets* are not a true measure of the natural discrimination of these dimensions, the findings may be an artefact of the design of the task. The baseline discriminability was calculated from the Baseline *conditions* in which the irrelevant dimension remains constant. This means that, in the Age *task*, age varied whilst sex remained constant. It may be – as I argue above – that age processing is based on the computation of featural information embedded within a configural context which is partially comprised of information relating to sex. If this is the case, holding the sex constant may be equivalent to maintaining the representation of a general configural template between trials in the Age *task*.

This might facilitate age processing in Baseline blocks since all that is necessary is to make judgements based on featural changes embedded within a largely unchanging configural template. In the Sex *task*, discrimination in Baseline blocks may be relatively more difficult because only the featural template remains constant, the underlying configural template still needs to be computed. It is assumed here that it is more resource-intensive to match configural templates given featural templates than to match featural templates given configural templates – that it is harder to perceive wholes given parts than to perceive parts given wholes.

Following this argument into the Orthogonal *conditions*, there is more interference in the Age *task* because it is necessary to process the features pertinent to age within a shifting configural context. This configural context must be processed to some extent in order to process the features embedded within it. In the Sex *task,* it is only necessary to process the configural context – one can more easily ignore the shifting features because the underlying configuration on which the decision is based is not dependent on them. This is why the asymmetrical effect occurs. As discussed above, both sex and age processing may rely upon both featural and configural processing, just to differing extents. This is why we still observe Garner interference of age on sex.

4.5.6 Stroop and redundancy interpretation

The Stroop and redundancy (value-level) effects seem to be broadly consistent with one another, and seem to be mitigated by task-difficulty. In general, there is more facilitation and less inhibition (in the positively and negatively correlated conditions respectively) in the harder *task* compared to the easier *task* - in which there was comparatively less facilitation and more inhibition respectively. The finding that redundancy costs were smaller than redundancy benefits suggests that correlated values facilitate performance more-so than uncorrelated values inhibit performance.

Greater and more beneficial interference in the harder task suggests that information from the values of the more discriminable dimension can be passively or actively recruited into the more difficult process. With regards to the benefits, this finding can be interpreted in terms of the SoPA – insofar as there is more time for the easier dimension to interfere with the harder dimension. The SoPA fails to account, however, for the redundancy benefits noted for the Sex *tasks* of the photographic *sets* (see figure 4.7) – although these effects were not significant.

We can better account for the results if we consider *why* the differences in processing speed emerge. This can be achieved by thinking in terms of featural and configural processing. Taking the photographic stimuli first, there were more benefits for the Sex *task* (even in the negatively correlated condition), and more costs in the Age *task*. Assuming that featural processing is faster than configural processing – it may also be that featural processing takes place in a more automatic fashion, being a bottom-up rather than top-down process (see Theeuwes, 2018). This would suggest that age perception may be a more automatic process than sex perception, and may proceed more from the bottom-up, whilst sex perception may proceed more so from the top-down – making use of pre-existing configural templates.

If there are redundancies between the values of sex and age to be exploited in processing, the faster, bottom-up featural processing which takes place more automatically may constrain and inform the activation of a top-down configural template associated with the correct sex judgement. This is why there are still benefits in the cost analysis, because even if the values are uncorrelated, the processing of features (carried out more automatically) still constrain the decision – they provide data upon which to base a sex decision. As for the benefits and costs in the Age *task* for the photographic *sets*, it may again be necessary to consider potential issues with the design of the experiment.

As discussed above, processing in the Age *task* may be facilitated in the Baseline blocks by the fact that the underlying configural templates associated with the sex of the target face remains constant between trials. In the positively and negatively correlated blocks, the irrelevant dimension varies, and redundancy effects are calculated as the difference between the Baseline blocks and the positively or negatively correlated blocks (for benefits and costs respectively). Given this, it is perhaps a surprise that there are any redundancy benefits in the Age *task* at all. One might expect that keeping the configural templates constant would lead to better performance than changing it – even if it's compatible with the required judgement.

For the Facegen *set* there was a large difference between *tasks* in terms of the Stroop effect, and a smaller difference between *sets* in the redundancy analyses. The Stroop finding may again be related to the unfamiliarity of the stimuli and the caricatured sex differences. This unfamiliarity may lead to a smaller benefit in the Sex *task* because the unfamiliar features were more difficult to recognise and thus facilitate sex processing. It may lead to a larger benefit in the Age *task* because easily matched configural templates (owed to larger differences between males and females in the Facegen *set*) facilitates featural processing. That is, being readily able to categorise the face according to sex by the activation of configural templates constrains the processing of the features embedded within those templates. The differences between *tasks* in the redundancy analyses were less severe for the Facegen *set*, which may simply be interpreted in terms of the SoPA – with less time for positive or negative interference in the easier *task* and more time in the harder task.

4.5.7 Differences between the dimension- and value-level findings

The most obvious explanation for the differences between the dimension- and value-level findings is that the former do not seem to be dependent on task-difficulty and the latter do. This may considered in terms of the findings of Boenke et al. (2009) that different temporal dynamics underpin Stroop and Garner effects – to the extent that Stroop effects are sensitive to intra-stimulus interference, and Garner effects to inter-stimulus interference.

This explains why the Garner interference was consistent across *sets* and the Stroop (and redundancy) interference was not. Value-level interference may be more sensitive to earlier processes, whilst the dimension-level interference may be sensitive to later processes – such as those involving the activation of semantic representations. The pattern of Garner effects may be immune to differences in sensory processing speed because Garner effects may be an index of the way that these dimensions interact at higher, more semantic and less visual levels of representation.

The uniformity of the Garner analysis results despite the differences in task difficulty suggests that the pattern of interference between the two dimensions is mitigated by something other than task-difficulty or stimulus dimensions. The making of an explicit Sex or Age judgement is an operation that is abstracted from a particular stimulus. This abstracted operation may be subject to the type of template-based hierarchical processing described above, with its own architecture – which accounts for why the Garner interference was uniform across *sets*.

Garner effects may therefore be reliant on memory – or pre-existing representational structures – whereas Stroop effects may not. This finding may provide a clue as to the stage in the perceptual process that the shared-signal between the levels of sex and age manifests. The redundancy effects that emerge between these dimensions may be more dependent upon shared visual properties of the stimulus, rather than shared underlying representational qualities. That is, Garner effects may occur later, or higher in the perceptual processing stream than Stroop and redundancy effects – at a stage that is more insulated from the specifics of the visual input.

4.6 Conclusions

I found bidirectional and asymmetrical Garner interference between the facial dimensions of sex and age, such that sex information affects age judgements more so than vice versa. The finding of bidirectional interference is bolstered by the Stroop and redundancy effects. The apparent insensitivity of the Garner effects to differences between stimulus *sets* in task-difficulty, and the apparent sensitivity of the value-level (Stroop and redundancy) effects to that difficulty, suggests a dissociation between the nature of the two types of interference – as has been suggested in the past (Boenke, 2009; Fitousi, 2020; Van Leeuwen & Bakker, 1995). This dissociation may be rooted in the level of perceptual processing at which the interference occurs.

Similarities across *sets* allow us to state with confidence that sex and age are not fully separable dimensions in visual face processing – it is not possible to selectively attend to one of these dimensions and filter the other one out of attention. Based on the Stroop and redundancy findings, we can also say that compatibility of values across dimensions facilitates performance to a greater degree than incompatibility inhibits it. The uniformity of the asymmetrical Garner interference suggests that sex information interferes with age judgements more so than vice versa.

I suggest that the utility of the speed-of-processing account (see Graham & LaBar, 2007; Melara & Mounts, 1993) can be increased if consideration is given to the mechanics underlying the discriminability between the dimensions of interest. I argue that sex perception is more reliant upon configural processing, and age perception is more reliant upon featural processing. Despite this differential reliance upon each type of process, I argue that each type of perception relies upon both. The overlap between these types of processing in sex and age perception may be part of what characterises their asymmetric perceptual independence.

I argue that consideration of the interaction between these two types of processing can provide a fuller explanation of the results – that accounting for the reasons underlying different speeds of processing can provide a richer context. This new account, which takes into consideration the speed-of-processing of different dimensions, but casts such differences in speed in terms of differential processing strategies (configural vs featural and top-down vs bottom-up) and the complexity of the necessary processing, I refer to as the depth-of-processing account (DoPA). The usage of the term “depth” captures the degree of complexity involved (Friedenberg, 2009). According to this account, sex processing is deeper (i.e. more complex) than age processing, which is “shallower” (i.e. simpler), and sex processing involves relatively more top-down processes and age processing involves relatively more bottom-up processing.

The analyses show that the use of different stimulus *sets* can result in significantly different experimental results, and so analyses based on a single *set* – particularly one chosen arbitrarily or without sufficient justification – should be interpreted with caution. The use of multiple *sets* in the current work has permitted us to draw more robust conclusions than if I had only used one – future research might consider a similar approach.

CHAPTER 5: Dimension and value level interference between faces: investigating integrality using the Eriksen flanker task

5.1 Abstract

In this chapter, I investigated whether the dimension and value interference between sex and age within faces in the Garner task extended to interference between faces in a flanker task. I presented target faces flanked by faces that were congruent or incongruent on both the task-relevant and task-irrelevant dimensions, in order to determine whether irrelevant information was automatically encoded from distractors and subsequently affected target processing. This task is thought to index interference at the response level. I found no evidence of a flanker effect, but evidence of a Stroop effect in an analysis of target properties. The lack of a flanker effect suggests that sex and age information from irrelevant faces is not integrated into sex or age judgements about a target, or that such information does not affect response level selection. Methodological issues which may have contributed to the lack of a flanker effect are discussed.

5.2 Introduction

When the visual sex and age of faces vary between serially presented single faces, these two social dimensions have been shown to interact with one another in visual processing. This interference has usually been found to be asymmetrical. Fitousi (2020) and Quinn and Macrae (2005) found that age information interfered with sex processing but not vice versa, whereas I found (Chapter 4, above) that each interfered with one another, but sex information generally interfered with age processing more so than vice versa. Such asymmetrical *interference processing* (Chen et al., 2022) is evidence of a failure of selective attention – meaning that one dimension cannot be processed without some processing of the other dimension taking place.
Such interference also appears to take place between the values of stimulus dimensions. It has been noted for instance that “young female” and “old male” faces are more efficiently processed than “old female” and “young male” faces (Fitousi, 2020; Kloth et al., 2015). The same is so for “Black male” and “Asian female” faces (Fitousi, 2020; Johnson et al., 2014) and “happy female” and “angry male” faces (Becker et al., 2007)

Whilst experiments concerned with such dimensions are unequivocal, and there is yet to emerge an accepted framework to fully explain the effects, they highlight the fact that complex and often asymmetric patterns of interaction characterise our selective attention towards socially relevant information in human faces. An open question is whether such interactions can occur between concurrently presented faces – whether value (e.g. Young and Female) or dimension level (e.g. age and sex) information from distractor faces can interfere with the processing of a target face.

Selective attention research at the level of the individual face is typically concerned with the processing of irrelevant aspects of the target stimulus. Research at the level of multiple, concurrent faces is concerned with the processing of irrelevant stimuli – a central issue in selective visual attention research (Bindemann, Burton & Jenkins, 2005). The former type of research is particularly useful at identifying and describing patterns of interference between dimensions and values in face processing, but it tells us less about the time course of such interference and the circumstances under which it comes about. Put simply, the former, intra-stimulus approach to investigating selective attention describes *what* processes occur - whilst the latter, inter-stimulus approach describes *when* and under which circumstances they occur.

Although it is clear that mechanisms of selective attention are operant in face processing, the stages of the perceptual process at which they tend to operate are less obvious. There is evidence to suggest that faces are a special class of object in terms of perceptual load and processing capacities, and need to be considered on their own terms. Faces have been shown to capture attention and distract from ongoing task-demands more so than other objects. They have been shown to slow visual search tasks when appearing within (Langton, Law, Burton & Schweinberger, 2008) and outside of (Lavie, Ro & Russell, 2003) the area in which the visual search task is taking place.

Young, Ellis Flude, McSweeney and Hay (1986) showed congruency effects in a name-classification task, by which performance was better in trials in which irrelevant faces of the same category (musician or politician) as the name to be categorised were presented, relative to those in which the faces were of a different category. Sato and Kawahara (2015) showed that distractor faces presented at short SOAs capture attention even when faces are completely irrelevant to a visual search.

Lavie, Ro & Russell (2003) found that increased perceptual load in a name-searching task led to the successful filtering of distractor non-face objects, but that – regardless of the perceptual load – distractor faces persisted in interfering. This was taken as support for the notion that face processing may be mandatory and that faces may be exceptions to the rule of perceptual load – being perceived and affecting task-relevant processing even when perceptual load was high. Going beyond this, Jenkins, Lavie and Driver (2003) argued for the existence of separable face and non-face object processing capacities.

They showed that congruency effects reduced in the presence of intact faces (compared to phase-shifted or inverted faces) in a task in which participants matched a face to a name – indicating that only increasing the size of a set of faces interfered with face-related processing. On the other hand, congruency effects in matching names to pictures of non-face objects were reduced by the presence of any form of distractor – including faces. This suggests not only that there are separable face and non-face object capacity limits, but that the former is only affected by other faces whilst the latter is affected by any object including faces. Qarooni et al. (2022) argued that the face-specific processing bottleneck must occur after detection, with face (and object) detection being a capacity free process.

Investigating the limits of face-specific processing capacities, Bindemann et al. (2005) conducted a variation on the name matching task carried out by Young et al. (1986). They paired either names or faces with distractors which could also be either names or faces and asked participants to make a sex or identity judgement on the target whilst ignoring the distractor. They found no congruency effects when both the target and the distractor were faces, despite finding that distractor faces led to larger congruency effects when paired with names than did other names. Names also operated as distractors on face target processing. This means that faces distracted from name processing and names distracted from face processing, but faces did not distract from face processing. This was taken as evidence that the face-specific processing capacity is such that only one face can be processed at a time.

Thoma and Lavie (2013) pointed out that previous perceptual load work on faces had involved an experimental task which was not strictly confined to face processing – typically involving a matching task (Lavie et al., 2003; Young et al., 1986) or a search task involving non-face objects (Sato & Kawahara, 2015). They conducted a visual search task with differing numbers of distractor faces. They asked participants to search for a famous face and then indicate whether it belonged to a politician or a pop-star. They found that RTs increased with the number of distractor faces (from one distractor to three), and found congruency effects (faster responses when the distractors were of the same category of famous person compared to when they were of the opposite) with both one and two but not three distractors.

This finding indicates that perceptual load impacts irrelevant face processing, but that the face-specific processing capacity may be able to process more than one face at a time – even to the extent of semantic information about distractors being processed. They argued that the conclusion reached by Bindemann et al. (2005) - that the processing capacity for faces is restricted to a single face - is likely to be incorrect. This argument is based on their own findings

– that face distractors are processed so long as there are less than three - and the fact that estimates for the perceptual capacity of all other types of objects are higher.

Lavie and Thoma (2014) conceded, however, that their own task was less challenging – and thus was of a lower perceptual load - than was that of Bindemann et al. (2005). They also argued that stimulus competition and reduced salience of the distractor relative to the target may have led to the demonstration of such austere face processing limitations for Bindemann et al. In short, task demands may modulate capacity limits for face processing, in certain cases reducing it so severely that only the target face can be processed at a given time.

Neumann and Schweinberger (2009) provided evidence for this by showing that attention to faces might be limited to a single object under conditions of high cognitive load, but not under low cognitive load. Their experiments involved the presentation of a target *within* a distractor – that is, for faces, a small target face superimposed into a large distractor face. Their high perceptual load task was an age judgement task, in which they found no interference effects. It should be noted, however, that all distractor faces in this experiment were rendered in greyscale whilst all targets were rendered in either red or blue – since a colour discrimination task was the low-perceptual load task. The targets and distractors were, in this sense, not naturalistic, and the added dimension of a constant difference in colour between target and distractor may have increased discriminability between them and enhanced the capacity for successful selective attention, free from interference.

It seems, then, that the processing of target faces can be subject to interference from distractor faces under some, but not all circumstances. Perceptual load seems to play an important role. The central outstanding question, then, regarding face-specific capacity limits, relates not to whether irrelevant facial distractors interfere with the processing of facial targets, but to what extent and under which circumstances. Taking task-difficulty as an index of perceptual load, it may be the case that more interference occurs for dimensions that are easier to process.

5.2.1 Eriksen flanker task

Questions such as these are often investigated using reaction time measures with concurrent targets and distractors, such as the flanker task, which was designed to determine the effect of distractor noise on target categorisation without the requirement of visual search (Eriksen & Eriksen, 1974). The flanker task has been used to determine the extent to which task-irrelevant, distractor information is automatically processed – on the assumption that some processing of irrelevant flankers occurs automatically, ultimately interfering with performance (Evans et al. 2011).

The flanker task involves the concurrent presentation of a single target - typically presented in the same central location between trials (to prevent the necessity of visual search) – with distractors, which are often either response compatible, neutral or response incompatible with the target (Eriksen & Eriksen, 1974). Typically, performance is facilitated when the distractors are compatible with the target, and is inhibited when the distractors are incompatible. Such interactions between the target and distractors are supposed to arise from the automatic processing of distractors.

The mechanism by which performance is thought to be affected is through competition at the level of response selection rather than sensory selection – with the level of sensory selection being that of feature extraction (Sanders & Lamers, 2002), However, response competition is supposed to arise as a result of physical similarities and differences between the targets and distractors, rather than in relation to their respective semantic content (Yeh & Eriksen, 1984). By this, then, the physical properties of the stimuli lead to interference at the behavioural level – sensory competition and selection leads to response competition and selection.

Based on this logic, distractor and target faces may interact with one another in processing - assuming that exemplars of given social categories have physical, sensory (that is, phenotypic) properties in common; for example, that there are common visual features between two females, or two young people. It may be argued that response competition may be affected when targets and distractors share or do not share socially relevant categories in common. Evidence for such effects from flanker experiments using faces remains inconclusive, although there are some methodological reasons why this may be the case.

In a series of flanker experiments, Fenske and Eastwood (2003) found that responses were faster when target faces were flanked by identical compared to different faces. They found that this effect was smaller when targets had negative compared to positive or neutral emotional expressions – although this effect disappeared when the features were scrambled. This led the researchers to the argument that attentional resources were modulated as a function of emotional expression, though these experiments were carried out with very simple schematic faces, and so the generalisability of these results to more realistic or photographic faces is difficult to determine.

Moser et al. (2008) investigated attentional biases towards valenced facial expressions in anxious populations using an Eriksen flanker task and ERP. They found no behavioural differences in the flanker task between anxious and non-anxious populations, but the ERP results highlighted the presence of an attentional bias towards faces with a negative expression in the anxious group, and a bias towards faces with a positive expression in the non-anxious group – in line with the hypothesis forwarded by Fenske and Eastwood (2003) that emotional expression may modulate attentional resources, and adding the dimension of individual differences.

In line with this, Ashley and Swick (2019) put forwards tentative evidence to suggest that PTSD patients may exhibit faster responses to angry expressions compared to controls in a flanker task. They showed that both control participants and those diagnosed with PTSD were slower to respond to emotional (angry or fearful) than to neutral stimuli, but that there were decreased RTs for the PTSD group for angry faces. This latter finding was revealed in post-hoc analyses with Bayesian modelling, and so the authors stressed the fact that further research into this area is necessary.

Importantly, Ashley and Swick (2019) did not control for the sex of their stimuli within stimulus arrays – sometimes showing a female target with male distractors or vice versa. This oversight of the dimension of sex, particularly when paired with emotional expression, may have confounded the results. As Becker et al. (2007) found that the dimensions of sex and emotional expression are not independently processed - and that the values of “angry” and “male”, and “happy” and "female" are compatible and lead to facilitated processing - any extent of the effect of emotional expression on RTs in Ashley and Swick's experiments may have been related to the sex of the target.

Soderberg and Sherman (2013) conducted an Implicit Association Test (IAT) with a flanker task, presenting either Black or White faces with either racially homogeneous or heterogeneous flanking faces to "non-Black" participants. They found that when Black faces were flanked by Black distractors there was an increased negative bias, and when White faces were flanked by White faces there was an increased positive bias. Both of these biases were found to be reduced when the faces were in a diverse context – that is, when their flankers were of a different race.

It should be noted in the above cited flanker tasks using faces as targets and distractors that interference seems to arise at the level of values and dimensions. That is, when we are trying to process one face, we may not process other faces in their entirety but may still encode some of the information about them. Information relating to a distractor face's emotional expression, race or sex may still impinge upon the processing of a target. This idea is consistent with the finding of Thoma, Ward and de Fockert (2016), who showed that not only faces, but horizontally misaligned face halves and polarity reversed face images acted similarly to faces when used as distractors in a flanker task, reducing the capacity to process other distractors.

The idea that distractor faces may not be fully processed, but that *features* of those faces may be processed, and thus have an impact on the processing of the target, is in keeping with the theory behind the flanker task. Current research using flanker tasks to investigate face processing has typically focused on emotional expression or race in isolation, however – given evidence for correlations between dimensional values in the shared signal hypothesis (Adams & Kleck, 2003; Becker et al., 2007; Kloth et al., 2015) it is necessary to explore how dimensional values interact. It has only been shown variously that such value level interference effects occur *within* stimuli, but not *between* them. The flanker task seems like a natural means of investigating this possibility. The current research employs a flanker task to investigate interference between the dimensions of sex and age, and at the level of their respective values.

5.2.2 Current Study

The current flanker experiment was designed to assess whether performance in a face categorisation task would be affected by flanking distractor faces – more specifically, whether social information from such faces would be automatically processed and thereby affect performance. participants were asked to make either a sex or an age judgement whilst both the sex and age of both the targets and distractors varied between trials. This means that, on a given trial, distractors could either be the same (congruent) or different (incongruent) on each the task-relevant and the task-irrelevant social dimension. If social information is automatically extracted from distractors, we would expect to find congruency effects – altered performance when the distractors are or are not congruent with the target on one or both dimensions. If the individual is able to selectively attend to the target without interference from distractors, we should expect no difference between conditions.

A third possibility is that there would only be congruency effects on the relevant but not the irrelevant dimension. This would be taken as evidence of affected competition at the response level and would suggest that the processing of targets and distractors was modulated by the adoption of a top-down attentional set which restricted attention to only those sensory features relevant to the target response. This outcome seems unlikely, given the documented integrality between sex and age, which is evidence of an incapacity to selectively attend to only one of these dimensions without processing the other.

As was pointed out by Kloth et al. (2015) and replicated by Fitousi (2020) and ourselves (Chapter 4), there appear to be congruency effects between the values of “male” and “old”, or “female” and “young” – meaning that responses are more efficient for either a sex or an age judgment when the target face displays one of these sets of values across dimensions as opposed to the alternative (“young male” or “old female”). Depending upon the extent to which information about distractor faces interferes with the processing of target faces, I expect that the processing of the value on the relevant dimension of the target face may be modulated by the value of the irrelevant dimension on the distractor faces. Such that, for instance, one may be faster to judge the target as “female” when the distractor is “young”.

Based on my previous experiment (chapter 4) and the literature reviewed above, I expect that social information about the sex and age of distractor faces will interfere with and modulate the processing of target faces. Furthermore, I expect that irrelevant sex information will interfere with age judgements to a greater extent than irrelevant age information will interfere with sex judgements – replicating the finding of asymmetric integrality between the two dimensions (chapter 4). An alternative hypothesis is based on perceptual load theory, by which we may find that more interference from distractor faces characterises the easier, more discriminable task – on the logic that more leftover processing capacity may operate on irrelevant distractors.

I will also conduct a Stroop analysis (Stroop, 1935; Fitousi, 2020; Kloth et al., 2015), in order to determine whether the presence of compatible values across dimensions of the target face facilitates performance – not accounting for distractor faces. Finally, I also expect to find compatibility effects such that the specific dimensional values of the distractor faces will affect the processing of targets, such that when the value on the irrelevant dimension of the distractor face is compatible with the value on the relevant dimension of the target face, performance will be facilitated.

5.3 Methods

Please refer to the *General Methods* section for more general information about this experiment.

5.3.1 Participants

A total of 380 participants took part in this experiment. 20 (5.68%) participants were excluded as outliers in terms of accuracy or RT, leaving 360 participants. 0.47% of individual trials were removed from the analysis (<200/>2500ms). Of the remaining participants, 353 (98.05%) reported demographic information. 180 participants (50.99%) reported their sex as male, 173 (49%) as female, and. The mean age of participants who provided age information was 27.25 (min = 19, max = 69). Participants were paid £0.88 (Median duration = 05:19). All procedures were granted ethical approval by the Research Ethics Committee of Bangor University's School of Psychology.

N	SEX	AGE	TOTAL
FACEGEN	60	59	119
MPI	61	60	121
LIFESPAN	60	60	120
TOTAL	181	179	360

Table 5.1: Number of participants per *task* and *set*

5.3.2 Stimuli

Three stimuli were presented in each trial, all drawn from the same *set*. There was a centrally presented target stimulus, and two peripherally presented distractor stimuli – one flanking each side of the horizontal axis of the target stimulus. To account for the cortical magnification factor (Daniel & Whitteridge, 1961) - by which foveal stimulation is of a higher resolution than peripheral stimulation– the surface area of each flanking distractor stimulus was 50% larger than the target stimulus. The presentation size of the configuration was set to be relative to the participants screen size on Gorilla; they were placed within a large central "zone" in Gorilla's Task Builder v1, which comprised around half of the available space. This ensured that the relative distance between the targets and flankers remained constant regardless of the device or screen size that the participant used to take part in the experiment.

5.3.3 Design and Procedure

In a mixed design there were two between-subjects variables, *task* (Age or Sex) and *set* (Facegen, MPI or Lifespan); and two within-subjects variables, *relevant* dimension (Congruent or Incongruent) and *irrelevant* dimension (Congruent or Incongruent). These within-subjects variables refer to the congruency of the distractor stimuli compared to the target stimuli. On each trial, the distractor flanking stimuli were either congruent or incongruent with the target, independently on both the *task*-relevant and *task*-irrelevant dimensions, yielding four conditions, “congruent-congruent”, “congruent-incongruent”, “incongruent-congruent” and “incongruent-incongruent” (see figure 5.1).

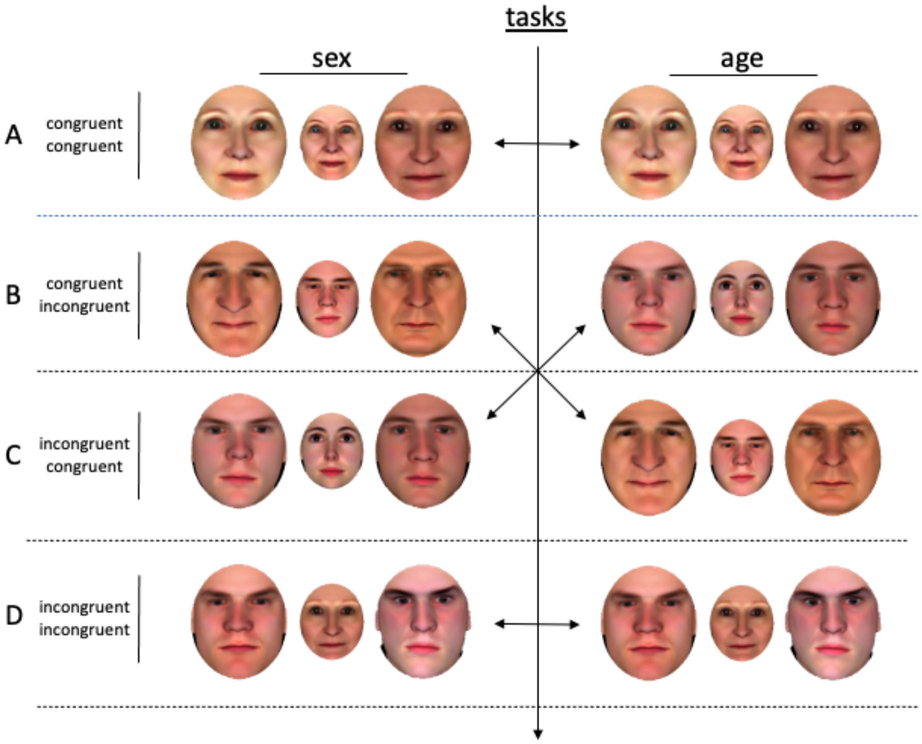


Figure 5.1: Example of each condition in the each task: A) relevant-congruent/irrelevant-congruent, B) relevant-incongruent/irrelevant-congruent, C) relevant-congruent/irrelevant-incongruent, D) relevant-incongruent/irrelevant-incongruent. These examples use the same configurations of target and flanker faces for each task, to illustrate that the condition that a configuration falls under depends entirely on the *task.*

The experimental phase consisted of four blocks, each containing 32 trials. There were 32 stimuli in each *set,* so each face was shown as a target once per block (four times in total across the experiment), once under each condition (see figure 5.1). The blocks and trials within blocks were presented in a random order. The only difference between *tasks* was in terms of the instructions informing participants to attend and respond to the relevant dimension (Sex or Age).

Participants were informed that each trial would consist of the presentation of three faces, and that their task was to respond as quickly and as accurately as possible to the face in the middle. The initial fixation cross was presented for 500ms and then the faces were presented for a maximum of 3000ms, or until a response was registered (see figure 5.2).

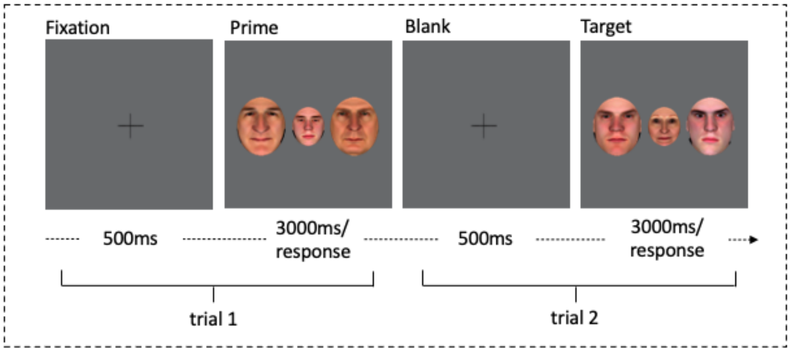


Figure 5.2: Example of two trials.

5.4 Results

5.4.1 Control Analysis

The control analysis was carried out in order to establish baseline performance for each *task* and *set* to inform the interpretation of the main analyses. I collapsed IES scores across levels of distractor relevance and used a two-way between subjects ANOVA with *task* and *set* as factors. This analysis revealed a significant effect of *task* ($F(1,354) = 10.02$, $p = 0.002$, $h^2 = 0.028$) such that performance was better in the Age (M = 826.66 (13.04)) than the Sex *task* (M = 882.54 (13.27)). There was also a significant main effect of *set* ($F(2,354) = 5.02$, $p = 0.007$, $h^2 = 0.027$). Fisher's LSD (45.25) indicated that performance was not significantly different between the Facegen (M = 833.99 (15.09)) and MPI (M = 836.4 (16.02)) *sets*, but performance was significantly better in both of these *sets* than the Lifespan *set* (M = 893.86 (17.22)).

These findings were qualified by a significant interaction between *task* and *set* ($F(2,354) = 20.6$, $p < 0.001$, $h^2 = 0.107$). Follow-up analyses showed significant differences for all three *sets*, such that performance was significantly better in the Sex compared to the Age *task* for the Facegen *set* ($t(117) = 3.4$, $p < 0.001$, $d = 0.62$; Sex M = 785.27 (20.59), Age M = 883.54 (20.29)), and was significantly better for the Age compared to the Sex *task* for both the MPI (t(118) = 5.92, p < 0.001, d = 1.07, Sex M = 919.43 (20.66), Age M = 752 (19.29)) and the Lifespan *set* (t(117) = 2.9, p = 0.004, d = 0.53; Sex M = 942.3 (22.61), Age M = 845.41 (24.61)).

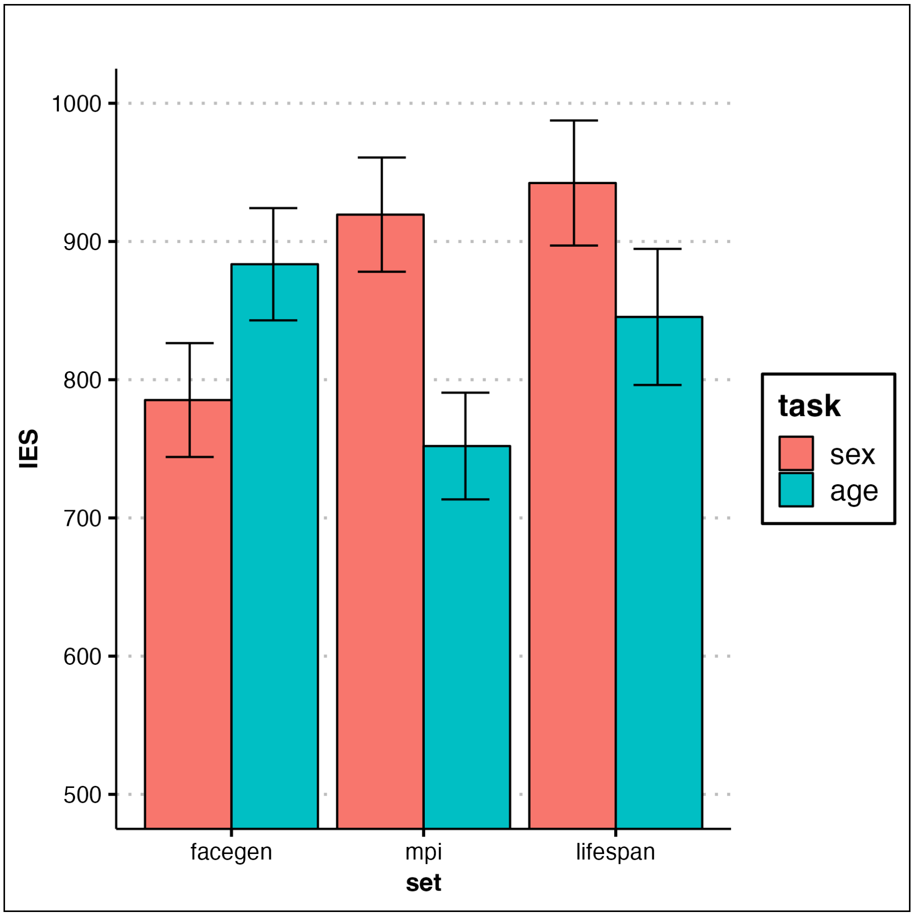


Figure 5.3: IES for each *set*, split by *task*.

5.4.2 Flanker Interference

The factorial design allowed assessing the independent effects of congruence between the distractors and the target, on the *task*-relevant dimension and the *task*-irrelevant dimension. To probe this effect, I generated a difference score by subtracting participants' performance scores (IES) on Congruent trials from their performance scores on Incongruent trials, separately for both the Relevant and the Irrelevant *dimensions.* This yielded two difference scores for each participant – one for each dimension (Relevant or Irrelevant). A positive difference score would provide evidence that Incongruent cues were associated with worse performance relative to Incongruent cues (flanker interference) and a negative difference score would suggest that Incongruent cues facilitated performance.

These difference scores were submitted to a 2x2x3 three-way mixed ANOVA with *relevance* (Relevant and Irrelevant) as a within-subjects factor, and *task* and *set* as between-subjects factors. Participants whose overall mean difference scores fell outwith 2.5 SDs of the mean difference score were excluded from the flanker interference analysis (n = 19).
This analysis showed no significant main effects or interactions (all F < 2, p > 0.13). Most importantly, there was no overall flanker interference effect, interpreted as a y-intercept significantly different from zero (F(1,336) = 0.01, p = 0.9; M = 0.22 (2.63)).

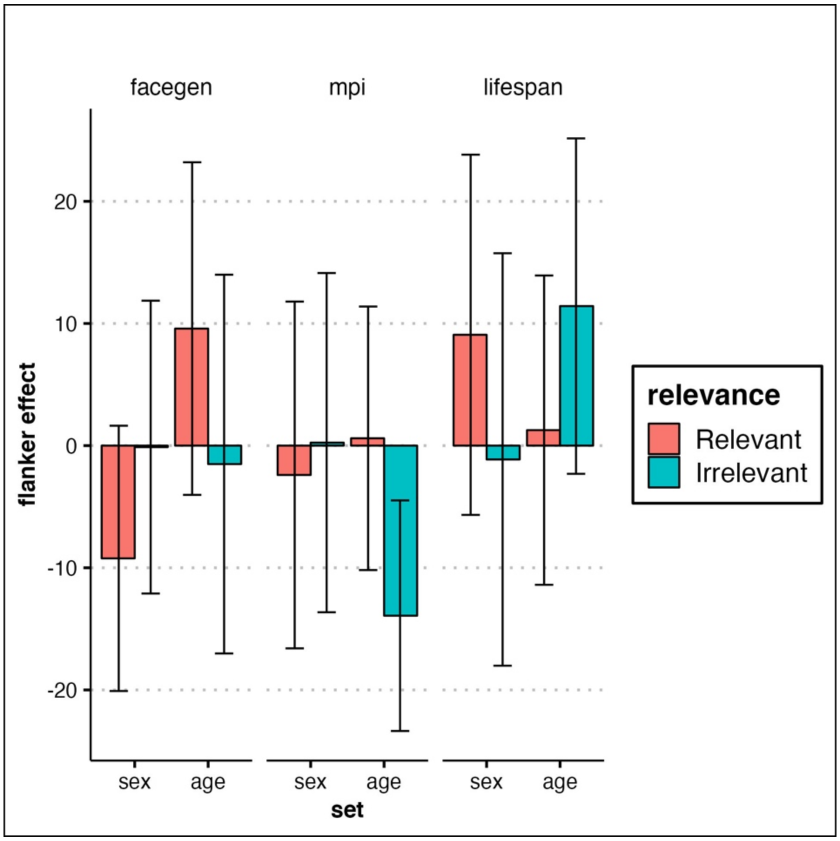


Figure 5.4: Flanker effects for each *task,* split by *relevance* for each *set* show no evidence of flanker effect.

5.4.3 Stroop Effects

Difference scores were calculated by subtracting the Compatible from the Incompatible configurations of dimensional values. Stroop effects and their possible modulation by *task* and by stimulus *set* were assessed by submitting difference scores to a 2x3 two-way between-subjects ANOVA with *task* and *set* as factors. Note that this analysis pertains only to the compatibility of sex and age cues from the target stimuli, not from the distractors. Prior to analysis, participants whose difference score (averaged across conditions) fell outwith 2.5 SDs of the mean difference score were excluded from the Stroop analysis (n = 15)

This analysis showed a significant, positive Stroop effect, interpreted as a y-intercept significantly different from zero (F(1,339) = 256.95, $p < 0.001$, $h^2 = 0.44$; M = 71.28 (4.91)), meaning that performance was significantly better for targets with Congruent compared to Incongruent combinations of levels of sex and age. This effect was modulated by *set* (F(2,339) = 8.24, $p < 0.001$, $h^2 = 0.046$; Facegen M = 93.71 (7.58); MPI = 68.06 (8.94); Lifespan M = 52.51 (8.55)). FLSD for *set* was calculated as 23.34, implying a significantly higher Stroop effect for the Facegen *set* compared to both the MPI and the Lifespan *set,* with no significant difference between the latter two. There was no main effect of *task* (F(1,339) = 0.57, p = 0.45) but there was a significant interaction between *task* and *set* (F(2,339) = 33.88 , $p < 0.001$, $h^2 = 0.167$).

Follow-up analyses showed significant differences for all three *sets* (see Figure 5.5). The Stroop effect was significantly larger in the Age compared to the Sex *task* for both the Facegen *set* (t(90) = 5.61, $p < 0.001$, d = 1.08; Sex M = 57.64 (7.41), Age M = 134.54 (11.52) and the Lifespan *set* (t(81) = 2.17, p = 0.03, d = 0.41; Sex M = 33.49 (14.31), Age M = 70.58 (6.98)) but the Stroop effect was significantly larger for the Sex *task* for the MPI *set* (t(81) = 5.89, $p < 0.001$, d = 1.1; Sex M = 116.19 (14.31), Age M = 22.34 (6.98)).

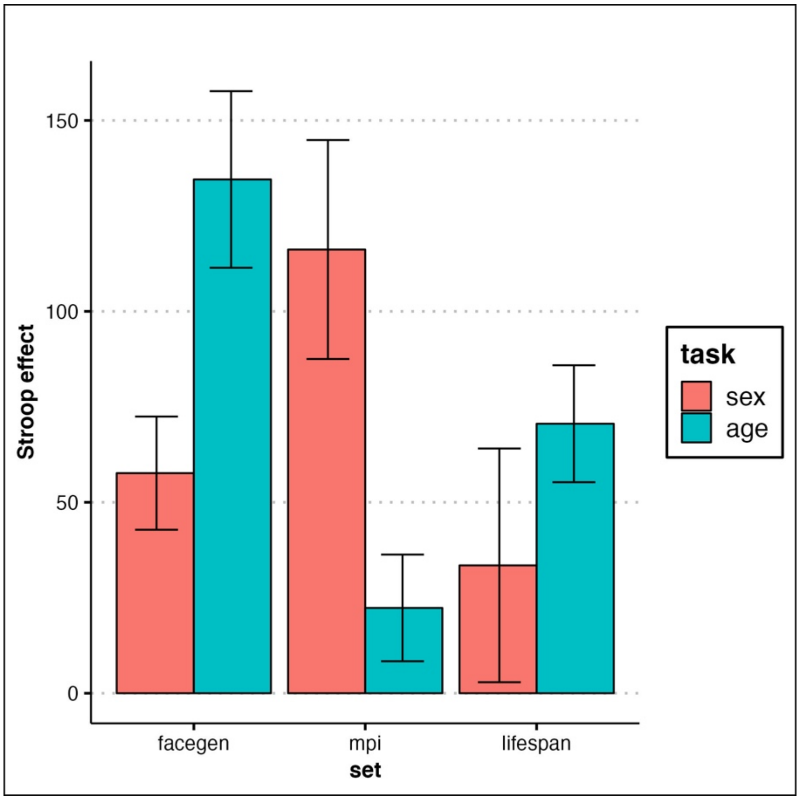


Figure 5.5: Stroop effect for each *set*, split by *task*. Positive values indicate Stroop effects.

5.4.5 Compatibility Effects

This analysis was similar to the Stroop effects analysis except in this case I was interested in how the irrelevant values of the distractor faces affected the processing of the relevant value of the target face. For example, in the Sex *task*, if the target face was Male and the distractor face was Old, this would constitute a Compatible trial, since “Male” and “Old” are expected to be compatible values (Kloth et al., 2015). An example of an Incompatible trial would be the same case except the distractor face was Young. Difference scores were calculated by subtracting the Compatible from the Incompatible conditions. A positive difference score would indicate better performance in the Compatible compared to the Incompatible condition, and would be interpreted as the presence of a compatibility effect.

As with the previous analysis, I assessed these effects and their possible modulation by *task* and by stimulus *set* by submitting difference scores to a 2x3 two-way mixed ANOVA with *task* as a within-subjects factor and *set* as a between-subjects factor.

These analyses showed no significant interactions or main effects (all F < 3.23, p > 0.07). Most importantly, there was a marginal overall compatibility effect, interpreted as a y-intercept significantly different from zero (F(1,345) = 3.23, p = 0.07; M = -4.83 (2.64)). This marginal but insignificant effect was negative, indicating a trend towards compatible values between distractors and targets inhibiting rather than facilitating performance.

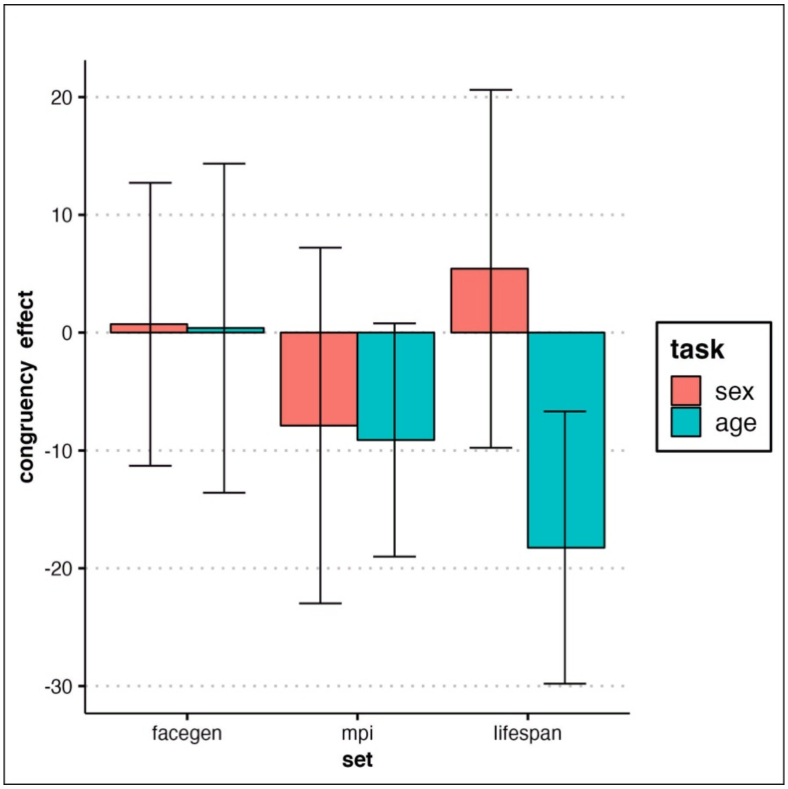


Figure 5.6: Compatibility effects for each *set*, split by *task*

5.5 Discussion

No evidence was found for flanker interference effects or value-level compatibility effects. There was evidence of Stroop effects in the expected direction, which constitutes a replication of my previous finding (chapter 4) and that of Kloth et al. (2015) and Fitousi (2020) - though it should be stressed at this stage that the Stroop finding relates to interference effects *within* the target face, not between target and distractor faces.

A close look at the differences between *sets* and *tasks* for the flanker and congruency analyses suggest that there may have been interference between target and distractor faces, but no such effects are consistently present in the current data to a statistically significant degree. These differences will first be briefly discussed before a discussion of the Stroop findings. I will then offer an overall interpretation of the findings before outlining some methodological issues and suggestions for future research.

5.5.1 Flanker and Congruency Findings

Different findings across the three stimulus *sets* makes an interpretation of the results difficult, but in the current case the utility of such an approach is on full display. A failure to find significant results for any of three different stimulus *sets* means that the failure to find an effect cannot readily be attributed to the stimuli, particularly given the finding of significant Stroop effects. Instead, there may have either been methodological and experimental design issues which resulted in a failure to find the compatibility effects between flankers and target - these are explored below. Alternatively, information about sex and age may not be automatically encoded from task-irrelevant distractor faces and hence have no measurable effect on the categorisation of target faces.

Although there were no significant flanker effects or modulations by *task* or *set*, it seems worth pointing out that there were some consistencies *within sets*. For example, in the Facegen *set,* for the Sex *task*, there was a slight negative effect of congruent, task-relevant information (Sex of the distractor the same as that of the target) and, in the Age *task*, a slight positive effect of congruent, task-relevant information (age of the distractor the same as that of the target). In contrast, the effects of task-irrelevant items were negligible for each *task.* This suggests that, for the Facegen *set*, when making a sex categorisation judgement and the sex of distractors was the same as that of the target, performance was inhibited, and when making an age judgement and the age of the distractors was the same as that of the target, performance was facilitated.

For the MPI *set*, there was a negative score on the Irrelevant dimension of the Age *task*, suggesting that age discrimination performance was facilitated when the sex of the distractors was different to that of the target. In the Lifespan *set*, incongruent sex information negatively affected performance in both tasks whilst incongruent age information affected neither task. The patterns of results were different for each *set*, but consistent within *sets* in terms of differences between levels of *relevance* and *task*. As for the congruency analysis, I again found no significant effects, but it is worth noting that – contrary to the hypothesis – the data trend in the negative direction for the two photographic *sets* (MPI and Lifespan). This means that congruency of dimensional values between distractors and targets appears to inhibit rather than facilitate performance.

Given these trends and patterns, I cannot say with confidence that the sex and age of distractor faces does not impinge upon target categorisation – these effects may just be small, or the current experiment may not be sufficiently designed to properly detect them. Whilst the findings are largely insignificant, the differences in trends between *sets* suggests that the specific stimulus properties may modulate the effect.

5.5.2 Stroop findings

There was a large, significant Stroop effect for each of the three *sets*. This was larger for the age compared to the sex *task* for the Facegen *set* and the Lifespan *set*, and larger for the sex compared to the age *task* for the MPI *set*. The Stroop effect was bigger for the harder task for the Facegen and MPI *sets*, but not for the Lifespan *set* – for which the Stroop effect was larger for the easier *task*. This effect for the former two *sets* is in keeping with the speed-of-processing account (SoPA) (e.g. Melara & Mounts, 1993) but not for the Lifespan *set*.

The Stroop effect shows the *difference* in performance between compatible and incompatible stimuli, so it does not indicate whether the effect is a result of facilitated processing of the former or inhibited processing of the latter. What *is* clear is that the Stroop effect is not characterised by response-level competition, since only the task-relevant response pattern is known to each participant, because task was manipulated between subjects. This implies that Stroop effects found here reflect redundancies in the stimulus structure across dimensions which facilitate or inhibit processing.

With that in mind, the finding that the larger Stroop effect occurs in the harder task seems at odds with perceptual load theory (Lavie & Tsal, 1994). According to this theory one might expect that the scope for interference might be greater in the easier task, since perceptual load is lower and therefore irrelevant stimuli may be processed to a greater extent owing to leftover processing capacity. Such an account would only explain the findings from the Lifespan *set*, but not the other two.

An alternative explanation for the other two *sets* would be that more task-irrelevant information may be recruited to make task-relevant decisions in harder, less discriminable tasks. Another way of saying this is that, in easier tasks, it might not be necessary to rely on cues from another dimension to make a decision. This is in keeping with the SoPA – by which the effects of the more difficult dimension on the easier dimension simply don't have time to manifest, the easier decision is made prior to any interference from the other dimension.

Of course, the perceptual load account and the differential discriminability account are mutually incompatible here – given that one offers an explanation for one *set,* the other for the other two. A clearer, mechanistic picture of what is driving the Stroop effect may emerge with more experimentation, and I return to interpreting the Stroop effect in the General Discussion (chapter 8) after reviewing the findings of this analysis from all of the experiments reported here.

5.5.3 Interpretation

The current results suggest that the sex and/or age of irrelevant distractor faces do not reliably affect categorisation judgements about central target faces. The current results also replicate previous findings (Fitousi, 2020; Kloth et al., 2015) of Stroop effects between specific levels of sex and age (Male and Old, Female and Young) *within* target faces. Whilst no significant flanker or compatibility effects between faces were found, flanker effects have been noted in the previous literature with respect to other dimensions (e.g. Ashley & Swick, 2019; Fenske & Eastwood, 2003; Moser et al., 2008), and so it is worth speculating on why we may not have found such effects here.

It has been suggested that facial expressions of emotion from distractor faces may interfere with the processing of target faces due to the adaptive and informatic value of emotional expressions – they communicate important social information relating to approach or avoid behaviour (Fenske & Eastwood, 2003). It has also been suggested that emotional expressions in experimental tasks do not bias attention in situations in which they are not task relevant (Lichtenstein-Vidne, Henik & Safadi, 2012; Victeur, Huguet & Silvert, 2019).

Sex and age information may have failed to modulate selective attention in the current experiment because of their relative lack of adaptive significance compared to emotional expression. Differential adaptive significance between dimensions is undoubtedly important and selective attention research may be confounded if such a factor is not taken into account. For example, one may reach conclusions about the underlying architecture of face perception based on asymmetric integrality between two dimensions that differ in terms of adaptive significance without accounting for that factor. The use of dimensions that are less emotionally and adaptively valenced may allow for more generalisation about the nature of the processing systems within which they operate – relatively unconfounded by the adaptive, informatic and emotional valence of different values of those dimensions.

The sex and age of the distractors may also have failed to affect selective attention because the distractors were consistently irrelevant to the task. As has been discussed above (chapter 4), failures of selective attention *within* the target face may be caused by overlapping featural and configural information processing operations occurring between the relevant and irrelevant dimensions, but such information does not overlap *between* different faces – particularly when some of those faces are task-irrelevant.

We can interpret the current results through the lens of Feature Integration Theory (Treisman & Gelade, 1980). The features of the distractor faces – although perhaps detected – are not integrated into perceptual wholes from which socially relevant information can be extracted because they are not attended to. Neither sufficient sensory nor semantic information about the dimensional values of distractor faces may emerge from the integration of the detected features, and thus those values do not interfere with the processing of the attended target face.

Alternatively, the grouping of the features relevant to a sex or age categorisation within a single face may make the location of those features easy to attend to, and their grouping in distractor faces easy to ignore. A social categorisation judgement may rely on the integration of features, and so – although there is evidence that faces themselves may be detected automatically (see Palermo & Rhodes, 2007) – categorisation may require the integration of features, a process that is not thought to be automatic and which is bound by conditions such as spatial grouping (Treisman & Gelade, 1980).

Such reasoning could potentially explain why interference processing has been shown to occur between faces where the dimensions of interest are emotional expression (e.g. Ashley & Swick, 2019) and race (Soderberg & Sherman, 2013). In these cases, it could be argued that certain features only need to be detected in order for a judgement to be made. For example, the detection of a smile, or of a certain skin-tone, could be argued to be sufficient to infer an expression or race judgement respectively. These dimensions are obviously more complex, and would require the integration of multiple features in order for a veridical judgement to be made in a real-world setting, but for systematically varying experimental stimuli, such features alone might suffice. That being said, I have suggested that age judgements can be carried out on the basis of features, and so – given the current logic – we should expect age cues from irrelevant faces to interfere more than sex cues.

5.5.4 Methodological issues

Were this experiment to be repeated, it would be advisable to make two changes to the design. The first would be to keep the two flankers identical. In the current experiment, the two flankers had the same demographic traits of sex and age, but were of different identities. If the two distractors on every trial had been of the same identity, this would have decreased the perceptual load on the participant and hence increased the likelihood of interference. The second change relates to the inclusion of a baseline condition. This would involve the inclusion of blocks in which there were no flankers, or in which the irrelevant dimension of the target face did not change between trials. Taking these in turn, the no-flankers condition would allow us to determine whether and to what extent the mere presence of faces as distractors affects performance.

Secondly, we could then see whether the effects of flankers facilitated or inhibited performance relative to a baseline. If performance was better in the no-flankers compared to the flanker condition, but we still failed to find differences relating to the congruency of dimensional values, we could reason that distractors do interfere with target processing, but below the level of the processing of processing sex and/or age.

The other type of baseline block, in which the irrelevant dimension of the target face remains constant, would allow us to better investigate the role of perceptual load – which is supposed to play a part in the extent to which face distractors interfere with target processing (e.g. Thoma & Lavie, 2014). In the Garner experiment reported above (Chapter 4) I found that performance was typically better in "Baseline" conditions (where the task-irrelevant dimension remained constant) compared to the orthogonal condition where both the relevant and the irrelevant dimensions varied. This suggests that the information-processing load is lower in the baseline compared to the orthogonal condition. All blocks in the current experiment were essentially orthogonal, in that the relevant and irrelevant dimensions of the target face constantly varied.

By designing the experiment in this way, and by the logic of perceptual load theory, we might find that distractor interference occurs to a larger extent in the baseline compared to the orthogonal blocks, since the perceptual load would be lower in baseline blocks compared to orthogonal blocks. Perceptual load may be too high when conducting the current tasks, and thus distractors may fail to be processed. The inclusion of either of these types of baseline conditions, in conjunction with using one identity instead of two for the flankers, may have led to the discovery of flanker and/or congruency effects or permitted us to make a stronger claim for the absence of such effects.

5.6 Conclusions

I found further evidence of interference *within* target faces, but no evidence of interference *between* distractor and target faces - or flanker effects. Whilst there are theoretical explanations for not finding such effects, consistencies within each *set* in the flanker and compatibility analyses and the methodological limitations of the current design discussed above mean that more research into the question is needed before strong conclusions can be reached.

CHAPTER 6: Investigating bottom-up and top-down interference on the perceptual integrality of sex and age using priming tasks

6.1 Abstract

Here, I used a face priming task and a word priming task to investigate the influence of irrelevant bottom-up and top-down information on the integrality of sex and age. Differences in the way that sex and age are influenced by these respective types of information provides clues as to how their representations are functionally organised in relation to one another. In the face priming experiments, there was a positive priming effect of congruent age information on sex judgements, but a negative priming effect of congruent sex information on sex judgements, and of congruent age and sex information on age judgements. This suggests that age information is irrelevant to sex judgements but sex information is relevant to age judgements – and potentially that sex information is more salient or difficult to ignore than age information and/or is prioritised in perceptual processing. In the word priming experiment, I found a positive priming effect for both relevant and irrelevant semantic information on target judgements. I also found a robust Stroop effect in each experiment. These effects provide further evidence of asymmetrical integration of sex and age. I discuss the modulation of these effects by stimulus set, and what this implies about the level of processing at which the perceptual integration occurs.

6.2 General Introduction

Thus far, I have investigated the interaction between visual sex and age information in human face perception at the level of the single face and at the level of multiple faces – exploring intra- and inter-stimulus information processing respectively. More specifically, I've been looking at our capacity to selectively attend to one of these sources of information whilst filtering out the other, and have found evidence of asymmetrical limitations of this capacity within individual faces. In the current chapter, I explore the sources of the attentional modulation operating between sex and age perception, assessing the respective contribution of bottom-up, stimulus-driven control; and top-down, psychologically-driven control. Below, I elucidate the problem of attentional modulation in more detail as well as the usefulness of the priming paradigm in exploring that concept.

I have chosen to investigate this crucial problem of attentional modulation using priming tasks. More specifically, I have carried out a face-priming experiment (using faces as primes) and a word priming experiment (using words as primes) to examine the role of bottom-up and top-down modulation respectively. Vision research concerned with this problem often involves visual search tasks (e.g. van Zoest & Donk, 2004; Wolfe, Butcher, Lee & Hyle, 2003) but such tasks are not always the most appropriate. Identifying different loci of attentional modulation in visual search tasks relies upon the differential distribution of attention across space, but here I am concerned with attentional modulation with respect to a single, centrally presented stimulus where the information underpinning the different dimensions of interest may be spatially overlapping – as in the experiments reported above (Chapters 4 and 5).

That is, I am interested in *analytic* selective attention (See Shalev & Algom, 2000), in the extraction of dimensional information from single objects – in this case, faces. It is likely that the information which people use to make sex and age judgements rely on attention to many of the same featural and configural properties of the face (Brown & Perrett, 1992; Rhodes, 2009) which overlap in space. The processes elucidated through visual search paradigms which involve the selection of objects or spatial regions within a display are therefore not relevant to the current investigation.

Like visual search paradigms, priming paradigms are concerned with attentional modulation, but seem more suitable for the analysis of analytical, intra-stimulus processing. Unlike visual search tasks, priming tasks typically involve the serial, rather than the parallel processing of distractors and targets. In visual search, targets and distractors are presented concurrently and are distributed throughout space. In priming experiments, the distractor (or “prime”) is typically presented alone before disappearing and being replaced by the target (or “probe”) – with both typically appearing at a predictable (often central) location.

In this, priming experiments are aimed at modulating attention to easily located targets presented in the absence of distractors rather than detecting and/or processing targets amongst concurrently presented distractors distributed through space. The question in priming experiments is not whether or how efficiently the target is detected and processed, but how the presentation of irrelevant information affects the processing of target information. They permit the presentation of physical and/or semantic information to the observer prior to the presentation of with a target.

The specific type of priming experiments described here measure “short-term priming” (Wentura & Rothermund, 2014). Within these types of experiments, a prime stimulus is presented which the participant is instructed to ignore, then – typically within a window of a few hundred milliseconds to a few seconds – a target is shown. The prime and the target are usually related in some way, sharing physical or semantic features in common. Priming effects are characterised by differences in categorisation performance when the prime and target are related compared to when they are not related.

The idea behind priming is that this physical or semantic information (which correspond to bottom-up and top-down controlling factors respectively) are provided at the prime phase will influence the observers processing of the target at the test phase, and thus affect their responses. The primed information is supposed to activate representations at some levels along the sensory-semantic axis of the perceptual processing hierarchy, rendering congruent responses to targets easier to initiate. In contrast, if the prime and target are incongruent, we might expect an inhibition of responses on account of competition between representations (Collins & Loftus, 1975).

For example, in the in the current experiments, I am concerned with the effects of primed information on the social categorisation of target faces. Such an effect – relatively better performance when the prime and probe stimuli are congruent compared to when they are incongruent – would be taken as evidence to suggest that the pre-activation of sensory or semantic level representations of a given dimension facilitates subsequent categorisation. In each task, participants were asked to make either a sex or an age judgement on a target face. In the face-priming task, participants were shown a prime face followed by a target face, in the word priming task participants were shown a word which relates to either the relevant or irrelevant dimensional information that characterises the upcoming target face – and the word is a valid cue to the target more often than not, engendering the expectation that the prime constitutes reliable semantic information for rendering a judgement.

In the face-priming experiment, I expect that some amount of the physical (sensory) information present in the prime face will be processed despite instructions for it to be ignored. If this information affects subsequent categorisation performance, it will be taken as evidence of the priming of socially relevant information from the bottom-up. In the word priming experiment, words referring to social categories are shown as primes. Since these words refer to social categories but are not physical representations of them, we can infer that any effect that their processing has on the subsequent categorisation based on physical aspects of the target must relate to the activation of semantic representations of the given social category - they therefore modulate attention from the top-down.

6.2.1 Interactive Activation and Competition (IAC) model of face priming effects

We can think about this bidirectional priming process in terms of a simple interactive activation and competition (IAC) model (Burton & Bruce, 1993; McClelland & Rumelhart, 1981). According to such a model, multiple dimensions would be represented as levels of a hierarchical structure, within which are nodes pertaining to each of the possible values of that dimension. From the bottom-up, perceptual information from sensory stimulation activates the node for which the psychological representation most closely fits the incoming data; from the top-down, expectations generated from the processing of semantic primes increases the resting rate of activation of those nodes – or “pre-activates” them (Perry, Lupker & Davis, 2008).

From the bottom-up – as explored through the face-priming experiment – one may, for example, see a young female prime. From this, we might expect that the nodes "Young" and "Female" become more active if the irrelevant social information from the prime is encoded. If we were carrying out the Sex task, and we were to allow for top-down modulation in our model, we might expect that the activation for the node "Female" would be stronger than for "Young" (see figure 6.1). From the top-down – as explored through the semantic-priming experiment – one may see the task-relevant priming word "FEMALE". This semantic activation, paired with the expectation that this will be a valid cue to the sex of the target, may pre-activate the representation of "Female", leading to the activation of that representation more easily, i.e. with less visual data needing to be processed before the categorisation can be made.

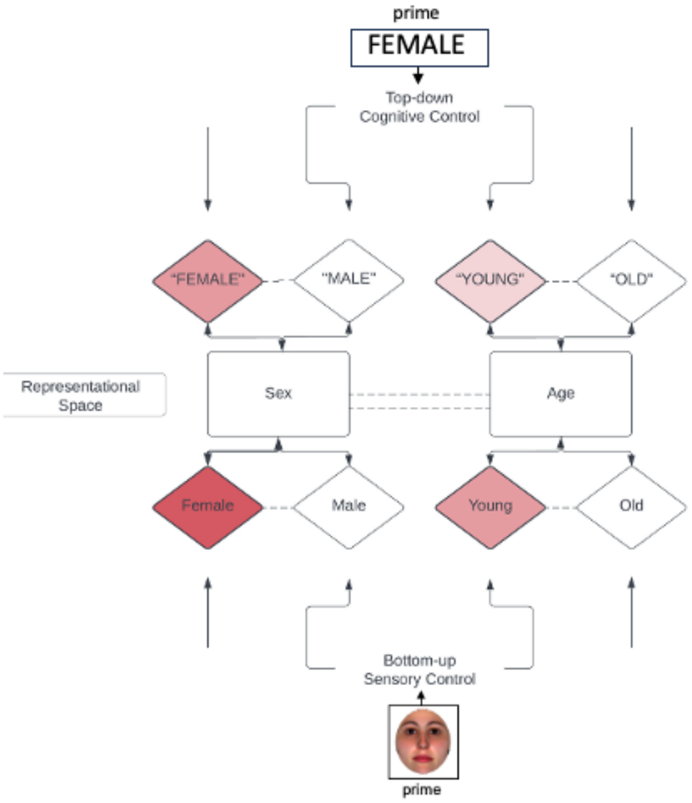


Figure 6.1: Simplified IAC model for Sex and Age processing in face perception. In an imaginary trial, the participant is presented with a Young Female prime in the Sex *task*. Here, we assume that the Young Female prime activates early visual representations of "Female" and "Young".

“Female” is more saturated than “Young” because of the possibility of task-relevance biasing activation through top-down modulation. From the top-down, semantic representations of these values are also activated to a lesser degree.

6.2.2 The current experiments

In the face-priming task, the congruency of the information between prime and target is balanced – that is, the prime shares the same sex and age category as the target on 25% of trials, just the same age on 25% of the trials, or just the same sex, or shares neither the same age nor the same sex. Faces, naturally, are multidimensional stimuli, and thus the face-priming prime contains information relating to both the task-relevant dimension and the task-irrelevant dimension. Generally speaking (although some caveats are discussed below) if information on both of these dimensions of the prime influences subsequent target categorisation, this will be taken as evidence of bottom-up modulation of social information – of automatic processing of sex and age information.

In the word priming task, the prime is a valid cue to the information about the target 75% of the time, creating an expectation that the cue can typically be relied upon to facilitate judgements. This expectation stands as a cognitive (higher-order) state which is biased towards one or another value. Single words are unidimensional, and thus the word priming prime is informative about either the task-relevant or the task-irrelevant dimension. Priming effects in the word priming task will be taken as evidence of top-down modulation of expectations pertaining to sex and age over subsequent social categorisation.

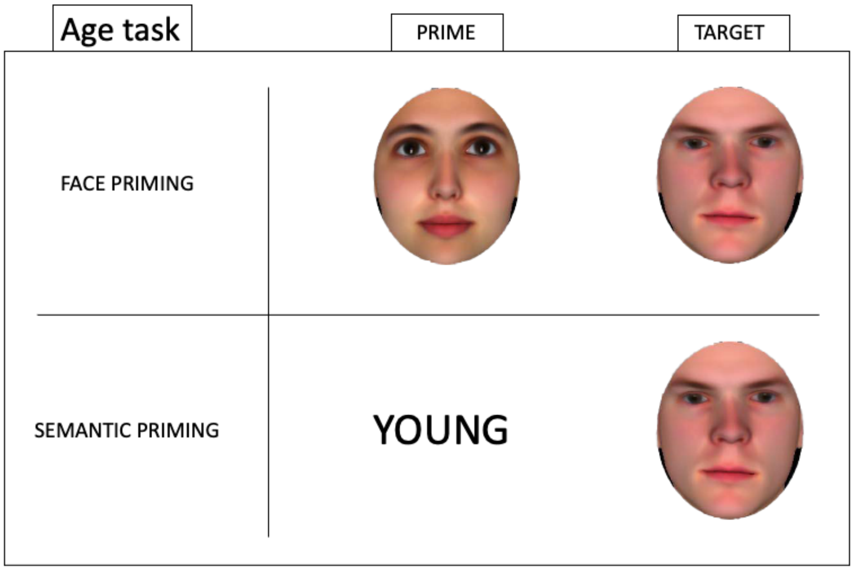


Figure 6.2: Simplified example trials of the Age *task* of both the face priming and word priming experiments. In the face priming example, the prime is congruent on the relevant dimension and incongruent on the irrelevant dimension (relevant-congruent/irrelevant-incongruent) and in the word priming example the prime is relevant and valid.

Although these two tasks are not exact counterparts, each can offer insights into the relative contribution of bottom-up and top-down modulating factors in sex and age perception. By being considered together they might uncover insights into the complex perceptual relationship between these two dimensions. In each of the experiments I expect to find priming effects when the dimensional values (either of the face or of the word in the face-priming and word priming experiments respectively) are congruent with the dimensional values of the target.

More specifically, I expect to find congruency effects – defined here as facilitation when the prime and the target are congruent on the primed dimension (or dimensions, in the face priming experiment) relative to when the prime and the target are incongruent. For example, in the word priming experiment, if the word-prime was “MALE” and the target face was a Young Male, we might expect to find facilitated performance compared to if the target was a Young Female. This would indicate that top-down modulation was at play, because the validity of the cue (75%) encourages participants to explicitly attend to the category specified by the word-prime. The null hypothesis is that there will be no priming effects. Such a result would imply efficient selection, i.e. the ability of participants to attend to the task-relevant aspects of the target face without also attending to related aspects of the task-irrelevant prime.

It has been noted above (Chapters 4 & 5) that sex and age are asymmetrically integral dimensions – that one cannot be processed without processing of the other taking place, but that sex information interfered with the processing of age information more than vice versa. Due to this, I expect to find asymmetrical priming between the sex and age judgement tasks in both experiments. I have suggested that sex may involve “deeper” processing than age, and that age may be integrated into sex processing. Based on this, I hypothesise that information relating to the sex of the target will have a greater effect upon performance than information relating to the age of the target – that age information will be easier to ignore and will have less of an effect on performance than sex information.

I have also suggested that sex processing may rely more upon configural templates which impinge upon the perceptual process from the top-down, whilst age processing may rely on relatively simpler featural templates – which, whilst still involving a top-down component, make age processing a more automatic and easier process driven more from the bottom-up relative to sex processing. Due to this hypothesis, I expect that bottom-up, physical cues to age might be extracted more easily and thus affect performance more in the face priming experiment, and that top-down, semantic cues to sex will facilitate processing more so than such cues to age in the word priming experiment.

6.3 Experiment 1: Face Priming

6.3.1 Introduction

The current experiment explores the effects of a face prime on the processing of a target face. It is assumed that any such effect will emerge as the result of the *automatic processing* of dimension- and value-information from the prime face. In order to explore whether and to what extent information about dimensional values is processed automatically, the current experiment tests the extent to which such information from irrelevant primes carries over into categorisation judgements made about target faces. Since faces are naturally multidimensional objects, the priming face will be congruent or incongruent with the target face in both sex and age. Because participants are asked to categorise the target according to either its sex or age, the prime will be congruent or incongruent along both a task-relevant and a task-irrelevant dimension.

If both sex and age are processed automatically, we should expect to find congruency effects, such that processing of the task-relevant dimension is affected by both dimensions of the prime face. On the other hand, if only the task-relevant dimension of the prime affects target processing, this could be explained by top-down processes such as the narrowing of the attentional window to dimension specific features of the face (Theeuwes, 2010), the gating of pre-attentive awareness (Kiefer, 2007) or the processing of the prime according to the same perceptual processing strategy adopted to process the target (Spruyt et al., 2007).

Thus, my main predictions hinge on whether the relationship between prime and test faces affects performance on the categorisation task, and does so equally or not for the task-irrelevant and task-relevant dimension of the prime. However, a further question is driven by my previous findings which indicate that sex and age are asymmetrically integral dimensions. Specifically, I found evidence for more extensive processing of sex in an age task than *vice versa* (chapter 4). How might this hypothetically asymmetric structure reveal itself in priming effects? This view predicts that, in an age task, the (mis)match between the sex of the prime face and the age of the target face would have a greater effect on performance than the opposite case (that is, the effects of a (mis)match between the age of the prime and target faces, during a sex task). Below, I briefly discuss the debate around automaticity of processing – or mandatory processing - with specific reference to the extraction of socially relevant information from faces.

6.3.1.1 The automatic processing of information

Cognitive models of attention often assume that distracting, irrelevant information may be processed automatically and in parallel, activating semantic representations of that information which may interfere with the processing of task-relevant information and representations. The activation of irrelevant representations must be inhibited in order to maintain efficient, coherent goal-oriented behaviour and thought (May, Kane & Hasher, 1995).

According to Theeuwes (1991, 1992, 2010) objects that are processed pre-attentively or "automatically" are perceived in a manner mainly controlled by bottom-up factors. There is, however, some general and broad disagreement as to what constitutes the criteria for automaticity of stimulus processing. A stimulus that is processed in an entirely automatic manner may be "mythical" according to Treisman, Vieira and Hayes (1992), who argued that we should perhaps think of automaticity more like a medical syndrome with a set of associated symptoms – the more symptoms of automaticity a stimulus has, the more automatised it's processing can be considered. One such set of symptoms was laid out by Palermo and Rhodes (2007), who argued that stimuli which can be processed rapidly, non-consciously, mandatorily and in a manner that is capacity-free – that is, requiring minimal perceptual resources – can be considered to be automatically processed.

6.3.1.2 Automatic face processing

According to Öhman (2002), faces are ideal candidates for this type of processing – although their research primarily focused on emotionally expressive faces. As far as whether faces "pop-out" – or can be found efficiently amongst other objects – there remains debate. Hershler and Hochstein (2005) reported "pop-out" effects for faces, but the validity and generalisability of these findings has been questioned (VanRullen, 2006). Wolfe and Horowitz (2004) argued that, since faces seem to be processed one at a time, they probably do not guide visual attention in a bottom-up manner.

On the other hand, faces can be processed and categorised as faces (rather than other types of objects) as early as 100ms (Liu, Harris & Kanwisher, 2002), which is around 80-100ms faster than for other objects (Pegna, Khateb, Michel & Landis, 2004). Non-conscious processing of faces has been observed for some people with acquired prosopagnosia (Tranel & Damasio, 1985) – who show higher levels of autonomic arousal for familiar compared to unfamiliar faces – and in healthy subjects it has been shown that information pertaining to identity and affect can be extracted from primes shown below the threshold of awareness (~17ms) (Stone and Valentine, 2004).

Mandatory processing of faces has been shown variously, with Suzuki and Cavanagh (1995) showing that performance in detecting the curvature of a line is negatively affected when such lines are in the configuration of a face. Young et al. (1986) found that the congruence of irrelevant face distractors to target names could facilitate or inhibit performance in matching the target name to a profession. Although the processing of irrelevant faces appears to be mandatory – and more resistant to manipulations of perceptual load than other objects (Lavie, Ro & Russell, 2003) – neuroimaging evidence suggests that responses to faces when they are task-irrelevant is attenuated (though not eliminated) compared to when they are task-relevant (e.g. O'Craven, Downing & Kanwisher, 1999).

Such mandatory processing may be restricted to faces themselves rather than dimensions of those faces. Murray, Machado and Knight (2011) carried out a similar task to that of Young et al. (1986), in which participants made speeded judgements about either the race or the gender of a name whilst ignoring a distractor face which was either congruent or incongruent with the task-relevant social dimension of the target name. They found congruency effects only under conditions of low perceptual load, leading them to the conclusion that these dimensions can be ignored, unlike faces themselves. They argued that attentional resources are required to encode non-identity aspects of the face, an argument against the mandatory processing of these dimensions.

This is in keeping with the evidence for capacity-free processing of faces, in which the picture is more complicated. It has been suggested that there are quite strict capacities to face processing (Bindemann et al., 2005; Thoma & Lavie, 2014) such that processing of faces is restricted to one, or very few faces. However, these capacities have been shown to be separable from the capacities for other objects (Lavie, Ro & Russell, 2003; Palermo & Rhodes, 2002) which suggests attentional resources that are specific to face perception.

Faces may, then, be processed automatically, but this may not extend to the dimensions of those faces. Faces may be processed holistically (Tanaka & Farah, 1993), but arguably not to the extent that information about specific facial dimensions is processed. Such an account would be in keeping with the holistic-to-analytic model proposed by Fitousi (2020). Fitousi used this model with specific regard to face dimensions, arguing that they are processed as integral initially but that their dimensions gradually become more separable. The same logic can be used to refer to faces themselves. It may be the case that the face is processed holistically and automatically at first, but only through the deployment of attention do the individual dimensions become discernible. At the automatic level of processing, dimensions may be integral with one another as well as integral with the face as a whole.

Some research has been directed at determining whether social information is extracted from faces automatically. It has previously been shown that categorisation according to facial dimensions requires top-down modulation, or the deployment of goal-directed action. Quinn and Macrae (2005) showed that mere exposure to social targets did not trigger categorisation, and that when exposure does lead to categorisation under dimension-specific encoding conditions, categorisation does not extend to task-irrelevant dimensions. They carried out three experiments: the first was a repetition priming experiment in which participants were either asked to passively view faces or make active sex categorisation judgements in the encoding phase, and then make sex categorisation judgements in the test phase which contained both repeated and novel faces.

They found faster responses to repeated compared to novel faces in the active but not the passive encoding condition. This repetition priming effect for actively encoded faces was taken as evidence that sex categorisation of the faces in the encoding phase did not occur automatically – and that social categorisation may be contingent on active processing goals. In their second experiment, they sought to determine whether this repetition priming effect occurred cross-dimensionally using sex and age judgements – whether actively encoding the sex of a face facilitated subsequent age judgements, for example.

They found that responses were faster when participants made the same judgement on faces that were repeated at both stages of the experiment (i.e. sex-sex or age-age) compared to making the same judgement but on a novel face – evidence of repetition priming when making an identical judgement. There was no evidence of this effect when the judgement changed between stages of the experiment (i.e. sex-age or age-sex) – evidence of a lack of cross-dimensional repetition priming. Interestingly, however, they found that categorisation of faces by sex was faster if it followed categorisation by age in the initial phase of the task, suggesting that irrelevant age information affected subsequent sex judgements but not vice-versa.

In a follow-up Garner speeded-classification experiment, Quinn and Macrae (2005) found asymmetric Garner interference between the two dimensions such that age information interfered with sex judgements but not vice-versa (the opposite pattern of asymmetry to that reported above (chapter 4)). In terms of the automaticity of sex and age processing, Quinn and Macrae's (2005) results suggest that task-irrelevant processing of facial dimensions may occur, but not to the point of categorisation occurring. Their finding of asymmetrical integrality suggests that, even when stimulus information which is relevant to multiple dimensions is encoded, this is not sufficient to engender categorisation. It is worth pointing out that greyscale "graphics files depicting faces" were used in these experiments, and so it can be assumed that the full wealth of stimulus information upon which one might ordinarily rely in processing social information – automatically or otherwise – is unlikely to have been available.

Reddy, Wilken and Koch (2004) presented evidence of sex discrimination from faces in the near absence of attention by conducting sex discrimination tasks under both single-task and dual-task conditions – finding insignificant differences in performance under each condition. This, however, cannot be taken as the automatic extraction of information about facial dimensions without attention, since the task always involved explicit sex discrimination. This finding may instead provide further evidence for separable attentional mechanisms specialised for faces. It may be that such attentional mechanisms were tuned towards sex discrimination in a top-down fashion, narrowing the attentional window towards features concerned with sex judgements through the adoption of a specific encoding strategy (Quinn & Macrae, 2005; Spruyt et al. 2007; Theeuwes, 2010)

It is not clear, then, whether face dimensions and their values are processed automatically and, if they are, to what extent. They may be processed, but not to the extent of interfering with ongoing task demands in another domain (such as the name-matching task of Murray et al. (2011)) or to the extent of triggering an explicit categorisation (Quinn & Macrae, 2005). The information that comprises the face must be processed if a face is to be registered, and information may be processed under passive encoding conditions but merely not to the point at which a categorical judgement necessarily occurs.

If faces are processed holistically (e.g. Tanaka & Farah, 1998), or if they are processed in a holistic-to-analytic fashion (e.g. Fitousi, 2020), then brief exposure to a face prime may not be sufficient for dimension and value level information to be extracted in a way that would interfere with the processing of a subsequently presented target. It may be, however, that different dimensions and values are automatically extracted at different rates – for instance, age processing may be more automatic than sex processing (Quinn & Macrae, 2005; Wiese et al., 2008, 2012). If age and sex information differ in terms of the automaticity of their processing, I should find asymmetrical interference effects in the current experiment.

6.3.2 Methods

Please refer to the *General Methods* section for more general information about this experiment.

6.3.2.1 Participants

A total of 197 participants took part in this experiment. 11 (5.58%) participants were excluded as outliers in terms of accuracy or RT, leaving 186 participants. 0.72% of individual trials were removed from the analysis (<200/>2500ms). Of the remaining participants, 170 (91.89%) reported demographic information. 87 (47%) reported their sex as male, 83 (44.86%) as female. The mean age of participants who provided age information was 31.74 (min = 19, max = 73). Participants were paid £1.08 (Median duration = 10:08). All procedures were granted ethical approval by the Research Ethics Committee of Bangor University's School of Psychology.

N	SEX	AGE	TOTAL
FACEGEN	31	31	62
MPI	32	32	64
LIFESPAN	30	30	60
TOTAL	93	93	183

Table 6.1: Number of participants per *task* and *set*

6.3.2.2 Design and Procedure

In a mixed design there were two between-subjects variables, *task* (judge face on either Age or Sex) and *set* (Facegen, MPI or Lifespan); and two within-subjects variables, *relevant* (Congruent or Incongruent) and *Irrelevant* (Congruent or Incongruent). The *relevant* and *irrelevant* variables refer to the congruence of the priming face with the target face on the *task*-relevant and *task*-irrelevant dimensions respectively. All of the experimental stimuli varied in both Sex (Female/Male) and Age (Old/Young), which means that all priming stimuli could be Congruent or Incongruent with the target stimuli along both of these dimensions simultaneously. This yields four trial types (see figure 6.3) pertaining to the possible combinations of Congruency across the *task-relevant* and *task-irrelevant* dimensions.

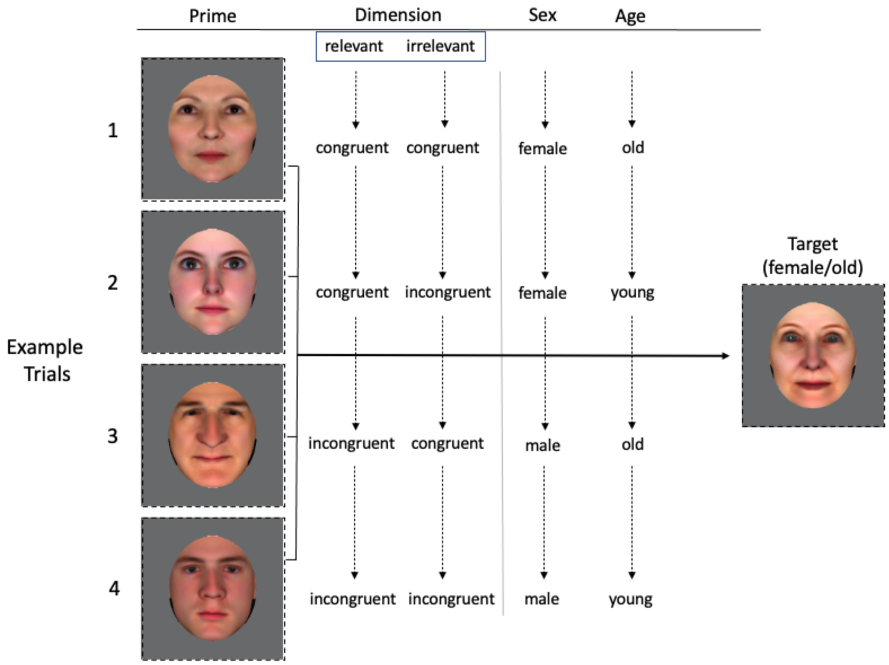


Figure 6.3: Examples of each trial type in the Sex *task*. Each of the prime examples in this case is relative to a target which is Female and Old.

Each trial consisted of the presentation of a fixation cross (400ms) followed by the face prime (200ms) followed by a blank screen (500ms) followed by the target stimulus (3000ms or until a response was registered) (see figure 6.4). There were four training trials, one for each trial type (see figure 6.3), followed by the test phase. There were four blocks, each containing 32 trials.

There were 32 stimuli in each *set,* and so each face was shown as a target once per block (four times in total across the experiment). In each block there were 8 trials for each trial-type (see figure 6.3, and so there were a total of 32 trials of each trial type throughout the experiment. For any given trial, the target face was paired with a priming face that met the specifications of the trial-type. An equal number of trials of each trial-type was presented in each block (n = 8 trials per trial-type per block) Both the prime faces and the target faces were of the same size and were drawn from the same set.

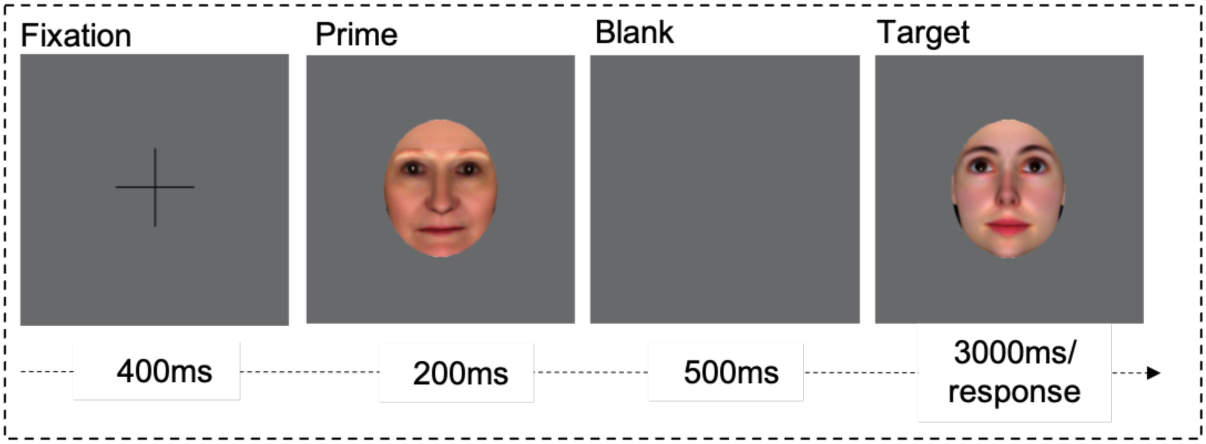


Figure 6.4: Example of a trial in the Age *task.* Here*,* the prime is Incongruent in terms of the *task*-relevant dimension and Congruent in terms of the *task*-irrelevant dimension (trial-type = Incongruent/Congruent). Since the target is a Young Female, the correct response would be “Young”.

6.3.3 Results

6.3.3.1 Control Analyses

A control analysis was carried out to establish baseline performance for each *task* and *set* to inform the interpretation of the main results. I collapsed IES scores on these factors across levels of distractor relevance and congruency and used a two-way between subjects ANOVA with *task* and *set* as factors. This analysis revealed a significant main effect of *task* ($F(1,180) = 12.71$ $p < 0.001$, $h^2 = 0.065$) such that performance was better in the Age *task* (M = 698.54 (13.58)) than the Sex *task* (M = 776.42 (19.27)). There was also a significant main effect of *set* ($F(2,180) =$

5.32, p = 0.005, h² = 0.055). Fisher's LSD (57.46) showed that performance was significantly different between the Facegen (M = 781.67 (16.64)) and the Lifespan (M = 781.69.16 (19.02)) *sets*, but not between either of these *sets* and the MPI *set* (M = 736.67 (24.82)).

These findings were qualified by a significant interaction between *task* and *set* (F(2,180) = 16.49, p < 0.001, h² = 0.15) (see figure 6.5). Follow-up analyses showed that performance was significantly better in the Sex compared to the Age *task* for the Facegen *set* (t(228) = 4.35, p < 0.001, d = 0.55; Sex M = 659.19 (26.06), Age M = 731.89 (18.93)), and was significantly better for the Age compared to the Sex *task* for both the MPI (t(226) = 10.27, p < 0.001, d = 1.28, Sex M = 849.55 (35), Age M = 623.8 (21.38)) and the Lifespan *set* (t(233) = 3.94, p < 0.001, d = 0.51; Sex M = 819.56 (27.92), Age M = 743.82 (24.38)).

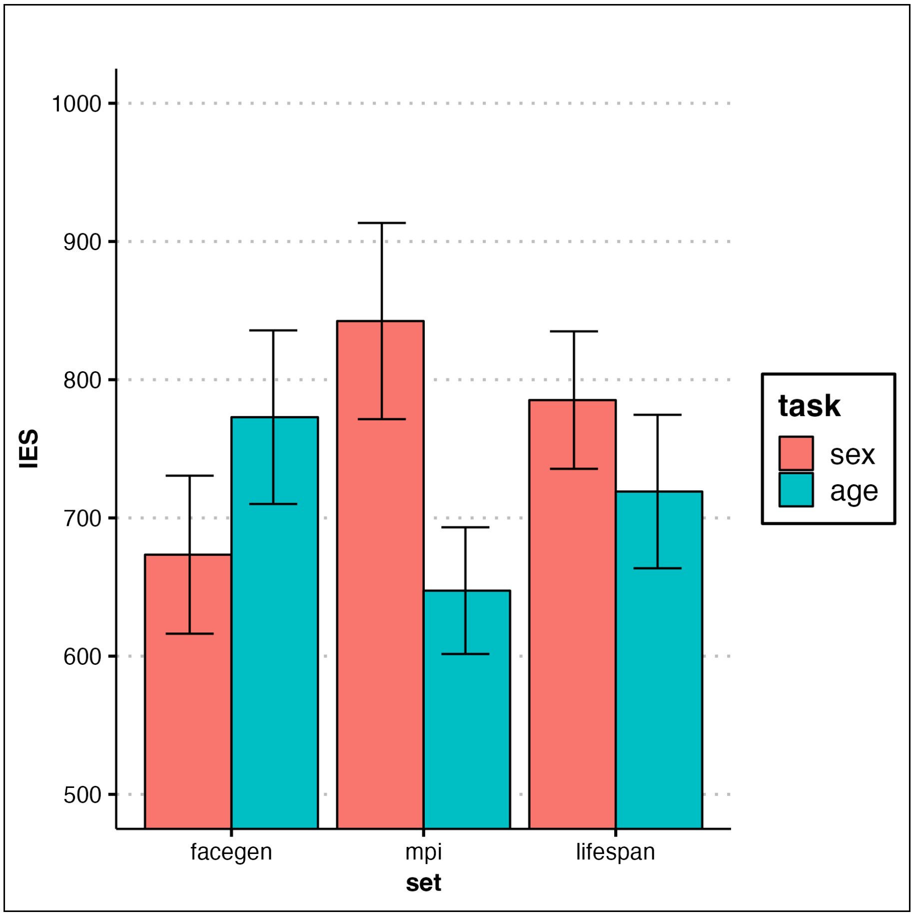


Figure 6.5: Control analyses for each *set*, split by *task,* designed to reveal overall differences in performance, independent of priming conditions.

6.3.3.2 Priming Effects

The presence of a priming effect was interpreted as a significant positive or negative difference score calculated by subtracting performance in the Congruent *condition* from performance in the Incongruent *condition.* This was performed separately for the *task*-relevant and the *task*-irrelevant dimensions, creating two difference scores per participant (Relevant and Irrelevant). A positive difference score would provide evidence of a priming benefit of Congruent relative to Incongruent cues (positive priming effect). A negative difference score would provide evidence of a cueing cost of Congruent relative to Incongruent cues (negative priming effect). Difference scores were submitted to a three-way mixed ANOVA with *relevance* as a within-subjects factor, and *task*, and *set* as between-subjects factors. Participants whose overall mean difference score fell outwith 2.5 SDs of the mean difference score were excluded from the priming effects analysis (n = 4).

This analysis showed a significant negative priming effect, interpreted as a y-intercept significantly different from zero (F(1,176) = 8.91 , $p < 0.003$, $h^2 = 0.026$; M = -7.88 (3.68)), which means that there was a cuing cost for Congruent relative to Incongruent cues. There was a significant effect of *task* (F(1,176) = 5.75, p = 0.02, $h^2 = 0.017$) such that the cuing cost was significantly larger in the Age (M = -14.2 (3.17)) compared to the Sex *task* (M = -1.56 (3.98)). There was a significant effect of *relevance* (F(1,176) = 10.36, $p < 0.002$, $h^2 = 0.027$) such that the cuing cost was larger for the Relevant (M = -15.93 (3.73)) compared to the Irrelevant dimension, in which there was no cuing cost (M = 0.16 (3.55)). There was also a significant interaction between *task* and *relevance* (F(1,176) = 4.72, p = 0.03, $h^2 = 0.012$).

Follow-up t-tests were significant for the Sex task (t(90) = 3.89, $p < 0.001$, d = 0.54) such that there were cuing costs for the Relevant dimension (M = -15.01 (5.39)) and a cuing benefit in the Irrelevant dimension (M = 11.89 (5.15), and were not significant for the Age task (t(90) = 0.74, p = 0.46, d = 0.11).

This interaction between *task* and *relevance* shows that, when the Sex of the prime is the same as that of the target, negative priming occurs regardless of the *task*. When the prime Age is the same as the Age of the target, negative priming occurs but only in the Age task (i.e. the

Relevant dimension of the Age *task*). Congruent Age between prime and target offers a cuing benefit in the Sex task (a positive priming effect). This means that congruent cues to Age facilitate Sex judgements but inhibit Age judgements, whereas congruent cues to Sex inhibit both types of judgement equally (see figure 6.6).

There was a marginal interaction between *task*, *set* and *relevance* (F(2,176) = 2.89, p = 0.058) which is characterised by greater cuing benefits for Irrelevant cues to the Sex *task* for the Facegen and MPI *sets* compared to the Lifespan *set* (see figure 6.7).

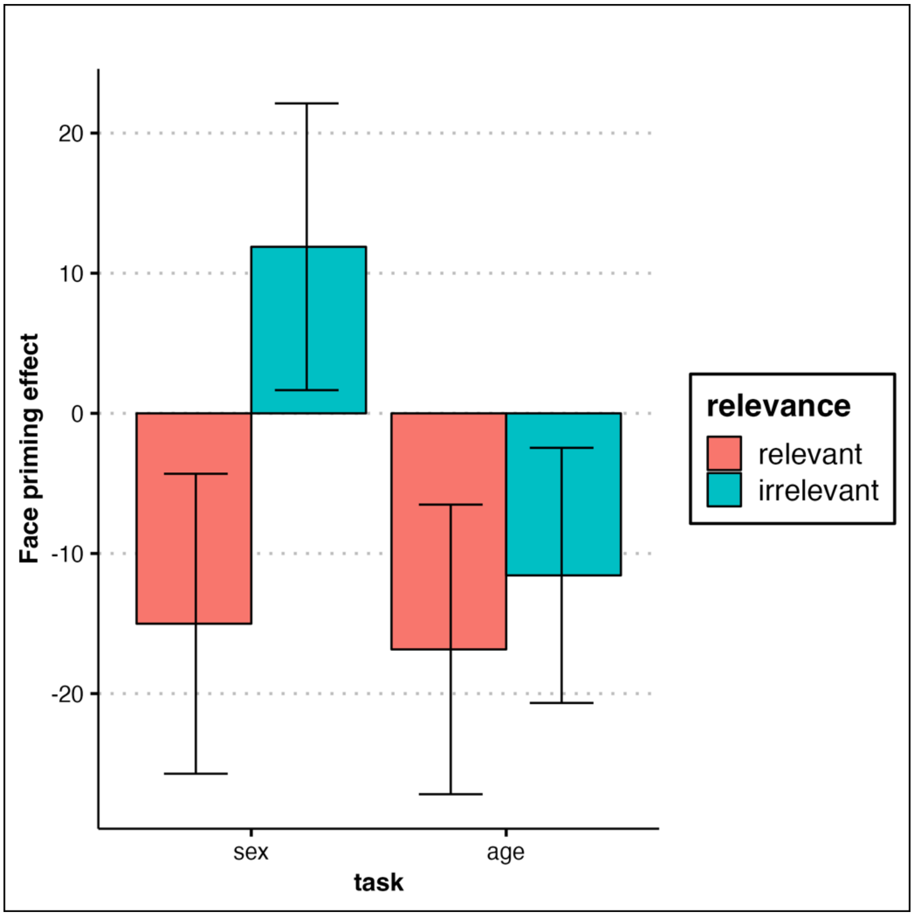


Figure 6.6: Effect for each *task*, split by *relevance*

It is worth noting the lack of a main effect of *set* (F(2,176) = 0.37, p = 0.69). This suggests that the asymmetrical priming effects noted above were broadly consistent between *sets* (see figure 6.7).

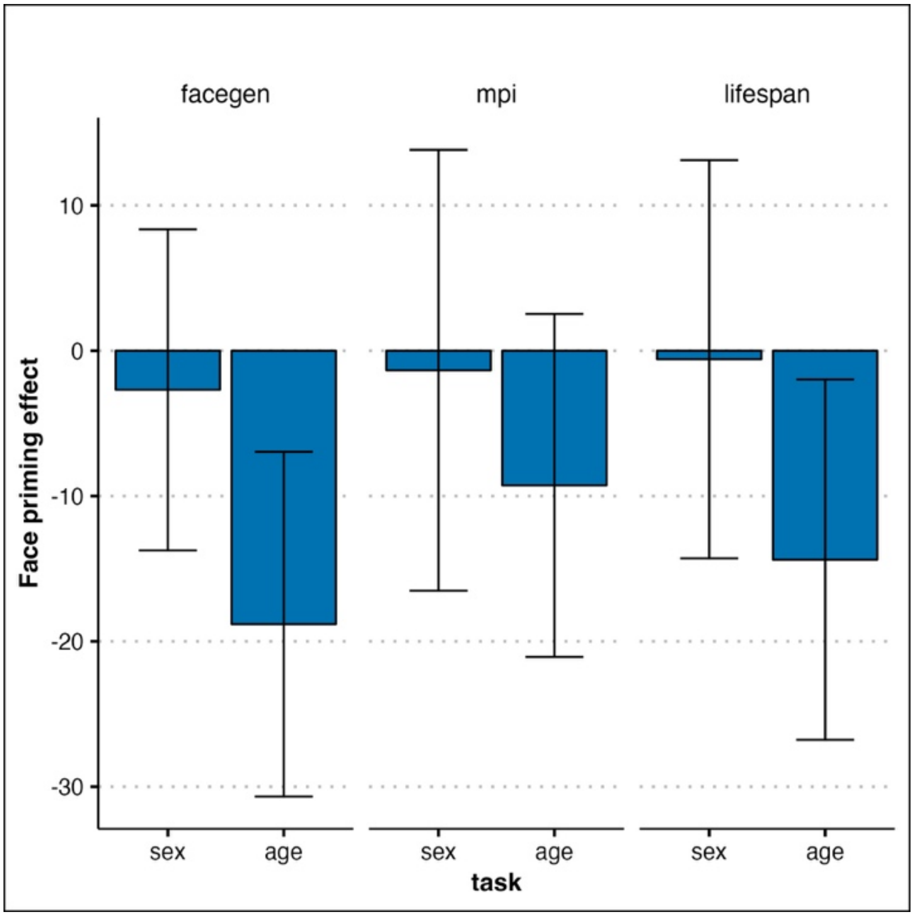


Figure 6.7: Asymmetrical priming effects for each *task* across *sets.*

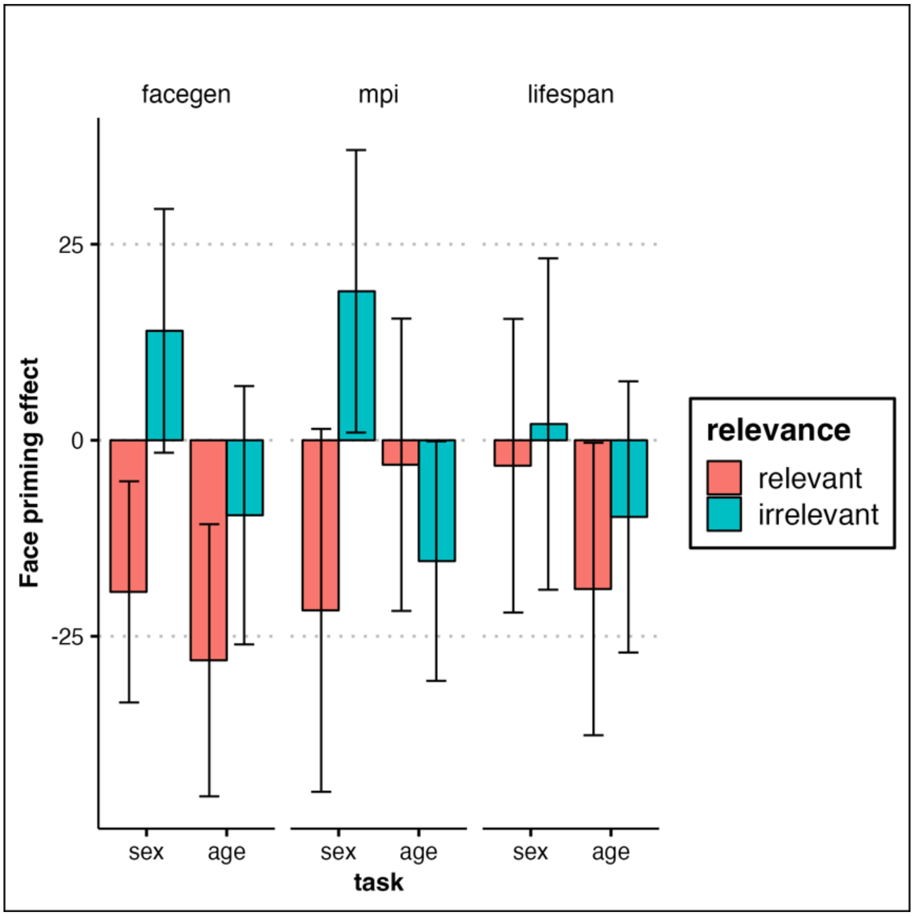


Figure 6.8: Marginal three-way interaction between *task, set* and *relevance* shows positive priming effect for Irrelevant Age cues on Sex task only – particularly for the Facegen and MPI *sets*.

6.3.3.3 Stroop Effects

I assessed Stroop effects and their possible modulation by *task* and by stimulus *set* by submitting difference scores to a 2x3 two-way mixed ANOVA with *task* as a within-subjects factor and *set* as a between-subjects factor. Note that this analysis pertains to performance in relation to the target stimuli only, and does not take into account the primes which preceded them. Participants whose difference scores fell outwith 2.5 SDs of the mean difference score were excluded from the Stroop effects analysis (n = 4).

The outcome of this analysis was a significant overall positive Stroop effect, interpreted as a y-intercept significantly different from zero ($F(1,176) = 115.30$, $p < 0.001$, $h^2 = 0.396$; $M = 56.16$

(5.68)). There was a significant main effect of *set* (F(2,176) = 3.74, p = 0.03, h² = 0.04). Fisher's LSD (27.16) showed that the Stroop effect was significantly larger in the Facegen (M = 76.04 (10.67)) *set* compared to the MPI (M = 46.91 (8.45)) and Lifespan *sets* (M = 45.81 (10.06)) but no difference between the MPI and Lifespan *sets*.

There was an interaction between *task* and *set* (F(2,176) = 12.91 p < 0.001, h² = 0.128). Follow-up analyses showed that, for the Facegen *set,* the Stroop effect was significantly larger in the Age compared to the Sex *task* (t(37) = 3.64, p < 0.001, d = 0.96; Sex M = 41.19 (7.68), Age M = 113.29 (18.25)). In contrast, In the MPI set, the Stroop effect was significantly larger in the Sex compared to the Age *task* (t(49) = 3.75, p < 0.001, d = 0.95; Sex M = 76.34 (13.24), Age M = 18.41 (7.95)). For the Lifespan *set* there was no significant difference between *tasks,* (t(41) = 0.84, p = 0.41; Sex M = 37.14 (18.13), Age M = 54.18 (9.32)) (see figure 6.9).

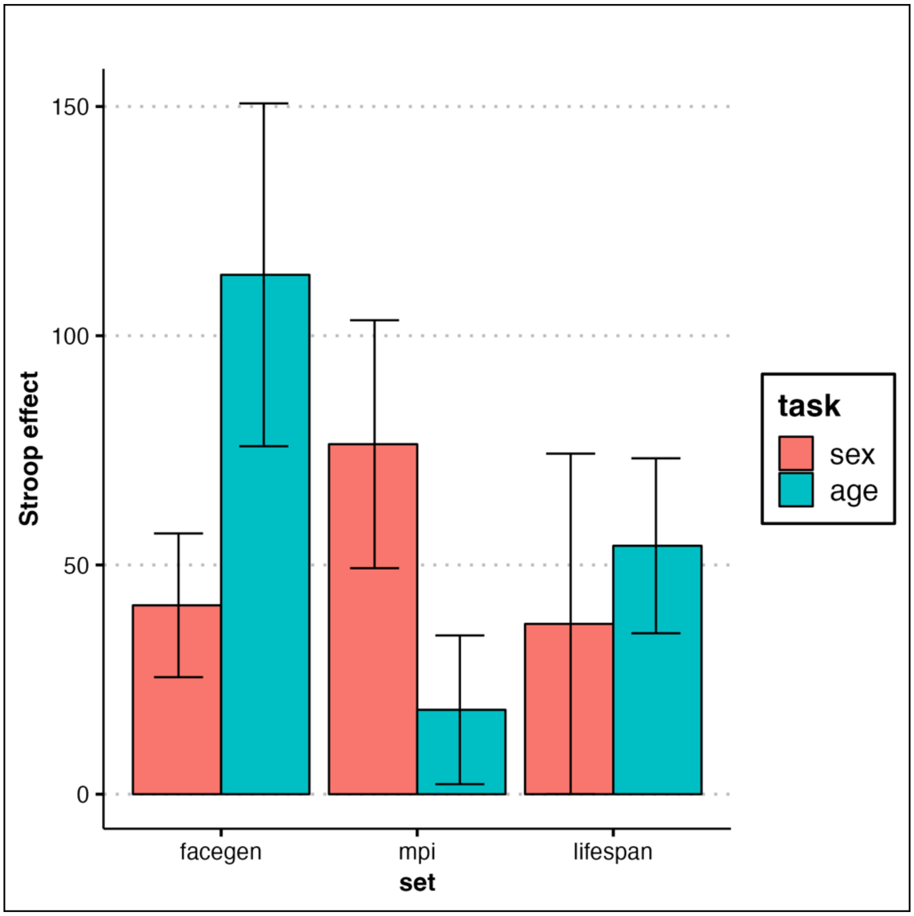


Figure 6.9: Stroop effect for each *set* split by *task*

6.3.4 Discussion

Contrary to the hypothesis, an overall, negative priming effect (Marcel, 1983; Tipper, 1985) was found when primes and targets shared a dimensional value that was relevant to the task. For example, when conducting the Sex (or Age) task, if the sex (or age) of the prime was the same as the sex (or age) of the target, there was a cuing cost. Interestingly, the direction of the priming effect of cues on the Irrelevant dimension varied by *task*, such that when the prime was the same sex as the target in the Age *task* there was a negative priming effect, but when the prime was the same age as the target in the Sex *task* there was a positive priming effect (see figure 6.6).

There was also a large, significant, Stroop effect for each of the three *sets* – showing that Old Male and Young Female faces are processed more efficiently than Young Male and Old Female faces. The extent of the Stroop effect differed between *tasks* as a function of *set*, being significantly larger for the Age than the Sex *task* in the Facegen *set* and vice-versa for the MPI *set*, with no significant difference between *tasks* for the Lifespan *set.* For the two *sets* in which there were significant differences between *tasks*, the greater Stroop effect was in the more difficult *task*. This trend was reversed for the Lifespan *set,* where the difference between *task* was the smallest and for which the difference in Stroop effects between *tasks* was not significant.

The priming findings provide evidence for the automatic processing of information relating to the dimensional values of both sex and age from task-irrelevant faces, and further support the hypothesis of asymmetric perceptual integrality of sex and age processing in human face perception. I found no evidence to support the hypothesis that age information may be automatically extracted at a faster rate than sex information. Below, I briefly describe the negative priming effect, and then offer two interpretations of the current results. I then review some methodological limitations and suggest directions for future research.

6.3.4.1 Negative Priming

The negative priming effect is defined as a performance cost found when a to-be-ignored “prime” becomes a subsequent to-be-attended “probe” (compared to controls in which stimuli are not repeated at the prime and probe phases). One interpretation of negative priming is that internal representation of the prime becomes associated with inhibition, thereby delaying the appropriate selection of that representation when it subsequently becomes task relevant. This is known as the distractor inhibition model, by which it has been suggested that distractor stimuli are initially activated and then deactivated, and that the top-down deactivation of distractor representations operates in parallel with the top-down activation of target representations (Frings, Wentura & Wühr, 2012).

In Tipper's (1985) experiments, trials consisted of a green and a red line drawing of different common objects superimposed over one another – with participants being asked to name the object in one colour and ignore the object in the other colour. RTs were higher for objects that had previously been ignored compared to novel objects. Negative priming occurred when the same exact stimulus was used at the prime and probe phases. The effect was also found when the prime and probe were semantically related. In Tipper's words, "this implies that ignored objects may receive analysis to a categorical level".

The negative priming effect is often characterised as resulting from interference at the level of response competition. Dalrymple-Alford and Budayr (1966), Neill (1977) and Tipper (1984) found in Stroop tasks that if the word ignored on trial N is the same as the colour naming response on trial N+1, performance latencies were longer. This was taken to suggest that the inhibition of the response to the distracting word on trial N carries over and impairs subsequent responses to the colour naming trial. In the current experiment, negative priming may have occurred if the prime exerted bottom-up modulation over social categorisation performance and inhibited the necessary target response. For example, if – in the Age *task* – a prime is a Young Male and then the target is a Young Female, the "Young" response may be inhibited, resulting in a negative priming effect compared to when the prime was Old and the target was Young.

Interestingly, however, in the case of the Age *task* there was a negative priming effect for both the relevant and the irrelevant dimension. This means that the interference cannot occur at the response level, since *task* was between-subjects and therefore the participant is unaware of the responses associated with the task-irrelevant dimension. According to Tipper (1985), if the prime and probe are well learned and meaningful to the subject then they may "achieve categorical internal representations that are beyond the level of specific physical features". Negative priming on the task-irrelevant dimension may mean that inhibition occurs at higher, more semantic levels – rather than at the lower, sensory levels of physical features or at the higher level of response selection.

As well as the negative priming effects, there was a positive priming effect – of age information in the Sex *task* (the irrelevant condition). Tipper (1985) suggested that positive priming indicates that the activation of the distracting element was above baseline (activated), whilst negative priming suggests that it was below baseline (inhibited). This would mean that the age information that resulted in positive priming did not need to be inhibited whilst making a sex judgement, whilst the sex information that resulted in negative priming did need to be inhibited whilst making an age judgement. The current finding therefore suggests that sex information needs to be inhibited in age judgements, but age information does not need to be inhibited in sex judgements.

This suggests that age information is not relevant to sex judgements, but that sex information is relevant to age judgements. Representations of sex (extracted automatically from the prime) must therefore be inhibited in making an age judgement whereas representations of age (also automatically extracted from the prime) need not be inhibited. Whilst we can interpret the negative priming effect for the Relevant dimension in each *task* in terms of increased response-competition, we cannot do the same for the effects on the Irrelevant dimension. This task-irrelevant interference must occur at a perceptual level. For example, in the Age *task*, it may be that irrelevant Sex information is extracted and takes priority over Age processing at either visual or semantic representational levels. This irrelevant sex information then competes and interferes with the relevant age information – leading to a declination in performance.

The findings provide further evidence of bi-directional, asymmetrical integrality between the dimensions of sex and age. Information about each dimension seems to be automatically extracted from a distractor (prime) face, and this information about each dimension seems to differentially affect decisions made about the other. The finding shows that sex information may be more relevant to age judgements than vice versa, which is why that information needs to be inhibited when making a judgement. Sensory (visual) or semantic level activity of sex information *competes* with that of age information, leading to negative priming. The activity of age information does not compete with that of sex representations, and instead facilitates judgements.

6.3.4.2 Interpretations

It is possible that participants may have been employing a top-down attentional strategy in order to process the task-relevant dimension of the target more efficiently, and that the prime stimuli were processed according to that strategy (Spruyt et al., 2007). However, if this were the case, we would only expect to see a priming effect in for the relevant dimension, since participants should only be employing a processing strategy to categorise the faces according to the task-relevant dimension, not the task-irrelevant dimension.

It may be that participants were unable to ignore the prime face because they had engaged an overall “face processing” strategy – that what appears as the automatic processing of irrelevant primes was actually constituted by an inability to selectively attend to target faces but not prime faces. Were this the case, however, we would expect symmetry between *tasks* and the Relevant and Irrelevant dimensions. Instead, what I found was asymmetric, mixed priming effects – although the majority were negative.

Although there were asymmetrical priming effects between the Irrelevant dimensions on each *task,* there were significant priming effects on each dimension of both tasks, meaning that both the task-relevant and the task-irrelevant dimensions of the irrelevant prime were processed. That is, despite instructions to ignore the prime, participants were unable to do so. In other words, the current result provides evidence that sex and age information modulate attention in a bottom-up manner - that information pertaining to these dimensions is processed automatically, and could be processed to the point of activating categorical representations.

The priming effects appear to be modulated less by task-difficulty than the Stroop effects (characterised by a marginal interaction between *relevance, task* and *set* (see figure 6.8)), which may suggest that the asymmetrical interference between sex and age which characterises the priming effects occurs at higher, more semantic and less visual levels of processing – although this cannot be said with confidence because of the marginal interaction. Below, I offer an interpretation for interference at lower, more sensory levels and higher, more semantic levels.

Some of the current findings can be interpreted in terms of the depth-of-processing account (DoPA) laid out above (see section 4.6), according to which sex perception involves relatively more complex and resource-intensive, top-down configural processing strategies, whilst age perception involves relatively simpler and less resource-intensive, bottom-up featural-processing strategies. Each type of process involves the matching of stimulus features to templates, or pre-existing semantic representations, but which are organised such that the templates relating to age perception operate at lower, more sensory levels whilst those relating to sex perception operate at higher, more semantic levels of processing.

The current results show that sex information interferes with age judgements but that age information facilitates sex judgements. This may be because of the depth of processing necessary for each type of visual processing. In the Sex *task*, when the prime provides congruent age information, the featural processing (relating to the age information) has already been done, meaning that all that needs to be done in order to make the sex judgement is carry out the necessary deeper processing. This renders the process more efficient than if there was no congruent age information to work from. Sex is considered to be relevant to age because sex information is processed at a deeper level of analysis – in this respect, sex processing underpins age processing, age information is embedded within sex information. Age is not relevant to sex because it is superficial in relation to it. Age does not affect the appearance of sex but sex affects the appearance of age.

Such thinking provides an explanation for why age information facilitates sex judgements, but fails to account for why sex information interferes with age judgements. One might expect that being given the necessary sex information would also facilitate age judgements – since the deeper, configural processing is already underway (having been extracted from the prime). One would think that this would make the relatively superficial and feature-driven age processing all the quicker, but this is not the case. The asymmetrical effect may therefore not be driven entirely by differences in visual processing strategy, but may emerge at higher, semantic levels of processing

According to such an interpretation, the task-relevant representation is inhibited because it is relevant to the judgement in the typical manner described in the negative priming literature (see

Tipper, 1985). Sex judgements are facilitated by age information in the same terms as described above: the age information extracted from the prime is not relevant, but it is helpful – some of the information processing that must take place on the target face has already been completed during the prime presentation.

The reason that sex representations need to be inhibited is because each sex ages differently, and so our representations of the age of each sex are likely to be different. Due to this, representations of age may be nested within representations of sex in the representational hierarchy (see figure 6.10). Due to this organisation, sex is relevant to age processing but not vice versa. This can be thought about in terms of parallel-contingent processing (see Johnstone & Downing, 2017), by which the two dimensions are processed in parallel, but the outputs of the processing of one dimension (sex) are passed onto and influence the other dimension (age), but not vice versa. According to such a scheme, age processing is contingent on sex processing.

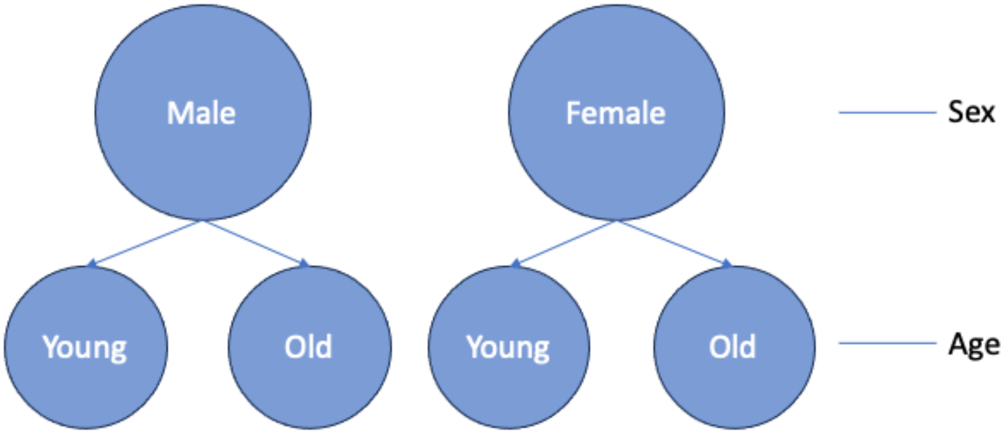


Figure 6.10: Simple hierarchy depicting the values of age nested within the values of sex in a parallel-contingent model

This means that, regardless of the discriminability of the stimulus features pertaining to the different types of information (sex or age), and the nature of the interference between compatible aspects of the stimuli for each *set*, the priming effect profile across *tasks* remains constant. This constitutes support of the hypothesis that representations of sex and age are asymmetrically hierarchically organised within the visual face perception system such that age processing is nested within sex processing. Within this context, sex can be used as a structural referent for age but not vice versa (see Ganel et al., 2005). This explains why age information can inform (facilitate) sex processing but sex information interferes with (inhibits) age processing.

6.4.3.4 Limitations

Although primes and targets were never the same stimulus within a given trial, and although the trials were organised within a block so that each target dimensional value and each condition of prime-target relationship was presented an equal number of times, trials (prime followed by target) were presented in a pseudorandom fashion. This means that there may have been occasions wherein the same response was necessitated multiple times in a row, or there were a number of stimuli in a row displaying the same dimensional value. It is likely, particularly given the results, that such incidents would affect processing, resulting for instance in repetition priming effects.

For example, it is conceivable that three or so Young Female targets might appear in a row, followed by some other configuration of values. As it stands, this simply constitutes noise in the data - particularly given the fact that the pattern of displays will have been different for each participant. A systematic investigation of inter-trial effects would have been useful. For example, if a participant was presented with three trials on which the irrelevant value of the prime remained the same (for example, they are carrying out the Sex *task* and the past three primes were Old), and then - on the fourth trial - the irrelevant value of the prime changes (to Young), there may be differential costs to performance compared to the reverse, or compared to when such a pattern of trials occurred in the opposite *task*. One might expect, given the current results, that there would be a bigger performance cost when doing the Age *task* compared to the Sex *task* - that changing the underlying Sex would disrupt Age processing more so than vice-versa.

An alternative would be to arrange the trials more carefully - so that the alternation between dimensional values was consistent across participants. Such designs might not only allow more information to be extracted from the data but might also reduce noise and facilitate the signal-to-noise ratio. Of course, such experiments may be very complex and require a huge amount of data to derive meaningful conclusions from.

The idea that what we have previously deployed our attention to might have a lingering effect on what we deploy our attention to next was described by Awh, Belopolsky and Theeuwes (2012) as “selection history” – an aspect of the perceptual process weighted equally with bottom-up and top-down modulate by Awh et al. (2012). The responses that we have previously made may impact future responses. I did not account for selection history in this experiment, but due to the pseudorandom presentation of trials which varied between participants, this is likely to constitute a source of noise rather than a confound.

6.3.5 Conclusions and future research

The current experiment suggests that sex and age information is automatically extracted from irrelevant and unfamiliar faces in face categorisation tasks. This information affects target categorisation performance even when it is irrelevant to the task. The finding of a negative priming effect of sex information on age judgements suggests that interference does not arise at the response selection level – since *task* was a between-subjects factor and participants didn’t know the response schedule for the irrelevant information.

The asymmetrical pattern of priming effects for irrelevant information between *tasks* (positive effect of age information in the Sex *task* and negative effect of sex in the Age *task*) provides further evidence of asymmetrical integrality between the dimensions of sex and age, and suggests that sex information is relevant to age judgements and age information is irrelevant to sex judgements. The current results also provide further tentative evidence for the idea that certain perceptual interference effects (such as priming effects) may occur at semantic rather than visual levels.

Although the precise mechanisms of the asymmetric interference are unclear, they may be explained in terms of differences in the depth (complexity) of processing required for each type of judgement and/or parallel contingent processing of representations of the values of the two dimensions. More research is needed in order to understand the mechanics and dynamics of the automatic, bottom-up extraction of socially relevant information from faces and how this information interacts in perceptual decision making. More specifically, it is necessary to disentangle visual and semantic level interference on the emergence of priming effects on response inhibition from those on psychological representation.

Potential future experiments may also be designed to account for the possibility of repetition priming and the effects of selection history. Future research should also attempt to discover whether socially-relevant information is extracted from faces in a task which does not itself involve face processing – in order to ensure that active top-down strategies involved in preparing participants for the act of face processing do not modulate the processing (or lack thereof) of the to-be-ignored face.

6.4 Experiment 2: Word priming

6.4.1 Introduction

Here, I explore the role of predictions in top-down attentional modulation – exploring its relation to social categories through stereotypes, the social cognitive phenomenon most commonly investigated using word priming paradigms. I then make a distinction between *congruency-based priming* and *validity-based priming* – and define the respective uses of these techniques in social cognitive research. I then make further definitions relating to "overtness" and "embeddedness" to refer to the different types of information expected to play a role in the priming effects explored in the current experiment.

As well as bottom-up visual cues, it has been shown that higher-order social cognitive states affect social perception through top-down modulation. For example, Anderson, Siegel, Bliss-Moreau and Feldman-Barrett (2011) found that faces which had been paired with negative social gossip dominated visual attention for longer in a binocular rivalry paradigm than did those paired with positive or neutral gossip. In the social perception/cognition literature, the top-down states investigated tend to be those that vary idiosyncratically between individuals - such as stereotypes, prejudices and attitudes. Such structures may be activated through priming, but not typically changed or affected by experimental manipulations.

These types of higher order states are also not “neutral” - that is, it is commonly assumed that some stereotypes, prejudices and attitudes are better or worse than others, and that a minimisation of stereotypy and prejudice in social categorisation is generally desirable. Relatively little social perception research is concerned with the role of ostensibly experimentally malleable and neutral top-down states, such as expectations about the upcoming category to be selected.

The idea that conscious or unconscious expectations, or predictions, affect perceptual processing has become so widely accepted that it characterises a number of currently influential theories of the mind, such as predictive coding (see Clark, 2013) and active inference (see Friston, 2006). Broadly speaking, such theories suggest that our senses are not sufficient to transform the noisy and ambiguous signals of the world on their own, and that perception is instead a process which involves using prior expectations or representations to predict - in a top-down manner - those incoming signals derived from the senses. One of the proposed mechanisms for this sort of process supposes that objects such as words are categorically represented in neural networks within the brain, and can thus exert top-down influence on incoming sensory signals (e.g. Macrae, Bodenhausen, Milne & Calvini, 1999; Harnad, 1986).

Encountering the word “CAT”, for instance, will be associated with the activation of the lexical representation of the category “cat” in a neural semantic network – the activation of the “cat” node (McClelland & Rumelhart, 1986). “Cat” would likely be closely connected with “dog” in such a network – since both are four-legged, commonly encountered, usually domesticated animals. Given this, the activation of “cat” would involve the partial activation of the representation of “dog” – leading an individual primed with “CAT” to be faster to respond to the target word, “DOG” than to the target word “PEACH”, for example (Rissman, Eliassen & Blumstein, 2003). Experimentally, such phenomena can be explored by investigating word priming effects – which can be said to have occurred when performance in processing a *target* stimulus is affected by the prior presentation of a semantically related *prime* (McNamara & Holbrook, 2003).

6.4.1.1 Congruency based and validity based word priming

Stereotypes can be thought of in terms of such semantic networks, being themselves networks of associated categories which contain the perceiver’s knowledge, beliefs and expectations about a given group (Macrae, Stangor & Hewstone, 1996). The activation of such a structure through priming is thought to make targets associated with a stereotype more accessible, leading to better performance in recognising targets when they are congruent with primes in terms of typically held beliefs. We can describe this type of priming as *congruency based word priming*. Within this method, the prime is more or less congruent with the target than another prime in terms such as semantic relatedness or stereotypy – making the method useful for investigating stereotypes or the structure of semantic memory.

A classic example of what I describe as a congruency based word priming effect comes from Fazio et al. (1986) who found that response times were enhanced between affectively congruent primes and targets (one would be quicker to respond to the word “HATE” after being primed with the word “UGLY”, for example, compared to when it was primed with the word “FLOWER”) – since “HATE” is more congruent with “UGLY” than it is with “FLOWER”. To put this in theoretical terms, “HATE” is more deeply connected in a semantic network with “UGLY” than with “FLOWER”. Extending this logic into the domain of social psychology, Banaji and Hardin (1996) found that participants were faster to characterise a pronoun as male or a female when the pronoun followed a prime which was stereotypically associated with that sex (e.g. “nurse”, “doctor”).

Similar findings have been noted with relation to racial stereotypes, even when the primes were masked. Wittenbrink, Judd and Park (1997) showed that White Americans reacted faster in a lexical decision task when implicitly primed with the word “WHITE” to a target word that had positive connotations (such as “INTELLIGENT” or “AMBITIOUS”), and when primed with the word “BLACK” to a target word that had negative connotations (such as “IGNORANT” or “VIOLENT”). In terms of sex judgements, Lemm, Dabady and Banaji (2005) used word cues which were either gender-stereotypic (e.g. “mechanic”, “hairdresser”) or words with gender-specific suffixes (e.g. “congressman”, “congresswoman”) to prime targets which were line drawings of people easily identifiable as male or female. They found priming effects (reduced RTs) for correct responses to male or female drawings following priming with gender-stereotypical job roles or words with gender-specific suffixes. They also found stronger priming effects for female than for male trials.

Another type of word priming that we can define is *validity based word priming* – where the prime is either a valid or an invalid cue to some meaningful aspect of the target. The logic of such designs is that valid primes may pre-activate or reduce the activation threshold of some critical aspect of the upcoming target, such that processing of the target is facilitated when the prime is valid, and interference and an inhibition of performance occurs when the prime is invalid. By this logic, A valid prime should generate an expectation about an upcoming stimulus relative to an invalid prime and should thereby be associated with better performance. Because validity priming offers a direct, rather than an indirect, cue to the target, it provides researchers a more direct measure of the attentional and perceptual processes that occur between the processing of the former and the latter – less affected by idiosyncratic differences in stereotypy, attitudes, prejudice or the structure of semantic memory.

What I refer to as *validity priming* has typically been used to investigate the effects of spatial attention on feature integration and detection, with Prinzmetal, Presti and Posner (1986) finding that valid primes to the location of an upcoming stimulus led to participants being less likely to make mistakes in reporting colour and shape information about that stimulus. Theeuwes, Kramer and Atchley (1999) found that detection performance in a visual search was hindered by invalid location priming. The method has also been used – although less frequently – to investigate visual object recognition and categorisation, including perceptual independence of stimulus dimensions.

An example of validity priming in social cognition is Johnstone and Downing's (2017) investigation into the perceptual independence of the dimensions of body weight and sex. They used word primes to either body weight or to sex in a sex categorisation or a weight categorisation task respectively - such that the prime always described the task-irrelevant dimensions of the target. They used an 80/20% ratio of valid to invalid prime so that the participant would develop the expectation that the prime was a reliable and veridical cue to a visual dimension of the target. They found an asymmetrical effect whereby the validity of weight cues did not affect the speed of sex judgements, but the validity of sex judgements affected weight judgements - invalid cues were associated with much higher latencies than valid cues. This finding shows that task-irrelevant semantic information can, but does not necessarily, affect subsequent judgements – and that expectations about upcoming stimuli affect their processing even when those expectations are ostensibly unrelated to the task at hand.

The result suggests that the representation of body shape is hierarchical, with a distribution of shapes associated with different weights clustered within each sex. This is a parallel-contingent interpretation, by which processing of the two dimensions occurs in parallel, but the computations of one dimension are contingent upon the other. The reason for this asymmetry was cast in terms of the proposed structure of the perceptual system, such that it may be organised in a way which parallels the real-world structure of the objects it perceives.

Male and female bodies gain and lose weight in different ways whereas people’s sex does not vary as a function of their weight. In this respect, Johnstone and Downing’s (2017) result can also be interpreted in terms of structural-reference (Ganel et al., 2004), such that sex is a structural referent for weight. An expectation about the sex of an upcoming target can therefore constrain and facilitate the processing of weight within that sex, whereas foreknowledge of a person’s weight cannot be used as a reference to compute their sex.

Because their task was between-subjects and since the prime was always to the task-irrelevant dimension, the finding of Johnston and Downing (2017) suggests that the priming stimulus did not merely facilitate the elicitation of a particular response – since participants could not be aware of the response mapping for the other dimension. That is, it appears that the prime activates representations at visual and/or semantic levels of processing which can be used to control subsequent behaviour in a top-down manner, rather than a simple response preparation. Their finding also suggests that representations of these dimensions are modality non-specific, that representations triggered by verbal cues can affect the subsequent processing of visual information.

That is, Johnstone and Downing (2017) showed that activation of a given socially relevant dimension could be achieved using “semantic” rather than “visual” cues – or top-down rather than bottom-up cues. Regardless of what this might say about domain-specificity, it is evidence of the fact that the nodes of perceptual hierarchies need not be activated in the visual processing domain in which their effects are observed – perceptual interference occurs between “types” or “modalities” of information.

6.4.1.2 The current experiment

In the current investigation, I extend this logic to the processing of sex and age information in faces – priming target faces with word primes. If, for example, we see the word “Male” as a prime during the Sex *task*, we might expect this to activate the semantic representation of maleness and - since there is an expectation that this prime is a meaningful cue to the upcoming social judgement - we should expect the activation of this representation to facilitate the judgement of a face as male.

I also expect to find Stroop compatibility such that Young Female, and Old Male faces are more efficiently processed than Old Female and Young Male faces (Fitousi, 2020; Kloth et al., 2015). It remains unclear whether Stroop effects such as those noted in the face-priming experiment reported above (and in chapters 4 and 5) emerge at the visual level or at the level of semantic, conceptual or psychological representation. In an attempt to address this question in the current experiment, the Stroop analysis will be carried out in such a way as to measure the effect of semantic cues on the Stroop effect.

It is important to draw a distinction between the types of information that will be measured in the priming and Stroop analyses respectively – with specific regard to task-irrelevant information. The difference between the sort of irrelevant information measured within the priming analysis and the Stroop analysis relates not only to the modality by which the information is received (as a word or a visual cue) but to the “overtness” of the information, and what I will refer to as its “embeddedness”. In the priming analysis, the irrelevant information that we are concerned with is overt and is not embedded; it is an unambiguous word presented in isolation which the participant is directed to attend to.

The same is not true of the irrelevant visual social information in the faces which may affect processing through its correlation to the particular task-relevant judgement being made. In this case, the information is covert – that is to say, attention is not directed to it – and it is embedded, that is to say, the information is presented within the same stimulus as is the task-relevant information to which the participant is directed to attend. It should be noted, however, that part of the current Stroop analysis involves measuring how overt, unembedded cues (semantic primes) modulate the Stroop effect.

We have here, then, two ways of exploring integrality – the first through overt and non-embedded primes, and the second through covert and embedded primes. An analysis of the respective contribution of these two types of information on task-relevant processing should help to elucidate the mechanics of the integrality of facial Sex and Age as well as to inform us about the role of top-down modulation in their processing.

This experiment involves presenting participants with a word which is a cue to either the relevant or the irrelevant dimensional value of the target (a valid cue 75% of the time). I predict validity effects, such that categorisation performance will be better following a valid compared to an invalid prime. I expect to find a symmetrical effect across tasks when the prime is relevant (i.e. a cue to the sex of the target in the sex categorisation task), but an asymmetrical effect when the prime is irrelevant (i.e. a cue to the age of the target in the sex categorisation task).

Given the findings that irrelevant sex information affects age judgements more than vice versa, (chapters 4 and 5), the hypothesis that sex and age differentially rely upon configural and featural processing (chapter 4), and the structural-reference (Ganel et al. 2004) and parallel-contingent (Johnstone & Downing, 2017) hypotheses, I expect that irrelevant primes will affect age judgements more than sex judgements. Irrelevant but valid (sex) cues to age judgements will pre-activate the structural referent for the subsequent judgement, and the dimension upon which the age judgement is contingent – constraining and directing the relatively less complex featural processing associated with the age judgement. That is, sex cues will constrain and inform age judgements because age information is embedded within sex information.

Thus, in age judgements, the configural template upon which the age judgement needs to be carried out will already be in place – primed by the irrelevant cue. In sex judgements, the featural template relating to the irrelevantly primed cue will be in place, but the underlying configural processing – which requires more attentional resources – will still need to be undertaken. Thus, irrelevant age primes to should still facilitate sex judgements, but not to the same extent as irrelevant sex primes should facilitate age judgements.

6.4.2 Methods

Please refer to the *General Methods* section for more general information about this experiment.

6.4.2.1 Participants

A total of 185 participants were recruited online through Prolific.co. 9 (4.86%) participants were excluded as outliers in terms of overall accuracy or RT, leaving 176 participants. 0.47% of individual trials were removed from the analysis (<200/>2500ms). Of the remaining participants, 174 (98.86%) reported demographic information. 85 participants (48.3%) reported their sex as male, 89 (50.56%) as female. The mean age of participants who provided age information was 37.57 (min = 18, max = 76). Participants were paid £0.94 (Median duration = 08:12). All procedures were granted ethical approval by the Research Ethics Committee of Bangor University's School of Psychology.

N	SEX	AGE	TOTAL
FACEGEN	29	30	59
MPI	28	30	58
LIFESPAN	29	30	59
TOTAL	86	90	176

Table 1: Number of participants per *task* and *set* after exclusions

6.4.2.2 Stimuli

The word cues were the words "MALE", "FEMALE", "YOUNG" and "OLD", written in upper-case. These were made using Microsoft PowerPoint, with the font *Calibri* and of size 96. The presentation size of the facial and prime stimuli was set to be relative to the participants screen size on Gorilla; they were placed within a large central "zone" in Gorilla's Task Builder v1, which comprised around half of the available space.

6.4.2.3 Design and Procedure

In a mixed design there were two between-subjects variables, *task* (Age or Sex) and *set* (Facegen, MPI or Lifespan); and two within-subjects variables, *relevance* (Relevant or Irrelevant) and *cue type* (Valid or Invalid). *Relevance* refers to whether the priming word was related to the task-relevant or task-irrelevant dimension. *Cue type* refers to whether the word cue was a valid or an invalid prime to the target. Participants were instructed to respond using the keyboard keys "F" and "J". The level of the *task*-relevant variable that each of these keys referred to was counterbalanced between participants, giving two response mappings by which to respond. Participants were randomly assigned to their *task*, *set* and response mapping using a Randomiser Node on Gorilla.

After consent was established, participants received specific instructions pertinent to the task, including instructions to attend to the *task*-relevant dimension. No mention whatsoever was made about the *task*-irrelevant dimension. Participants were instructed to respond as quickly and accurately as possible. There were four training trials, two with Relevant cues, two with Irrelevant cues; all of the training cues were Valid. After that was the testing phase, in which each trial consisted of the presentation of a fixation cross (1000ms) followed by the word prime (750ms) followed by the target stimulus (3000ms or until a response was registered; see Figure 6.2). The instructions and experiment took place against a grey background.

Each block consisted of 32 trials. There were 32 facial stimuli in each *set;* each stimulus was shown as a target once per block. The experiment consisted of four blocks, two "Relevant" and two "Irrelevant" blocks, the presentation order of which was randomised between participants. 25% of the cues in each block were Invalid (8 out of 32); the remaining 75% of cues were Valid. Within each *set*, there are four subsets (Young Male, Old Male, Young Female, Old Male). Two stimuli from each of these subsets was selected to be paired with an Invalid prime word in each block – reaching the number of 8 Invalid primes per block (see figure 6.11). This resulted in each of the 32 stimuli being paired with an Invalid prime one time over the course of the experiment.

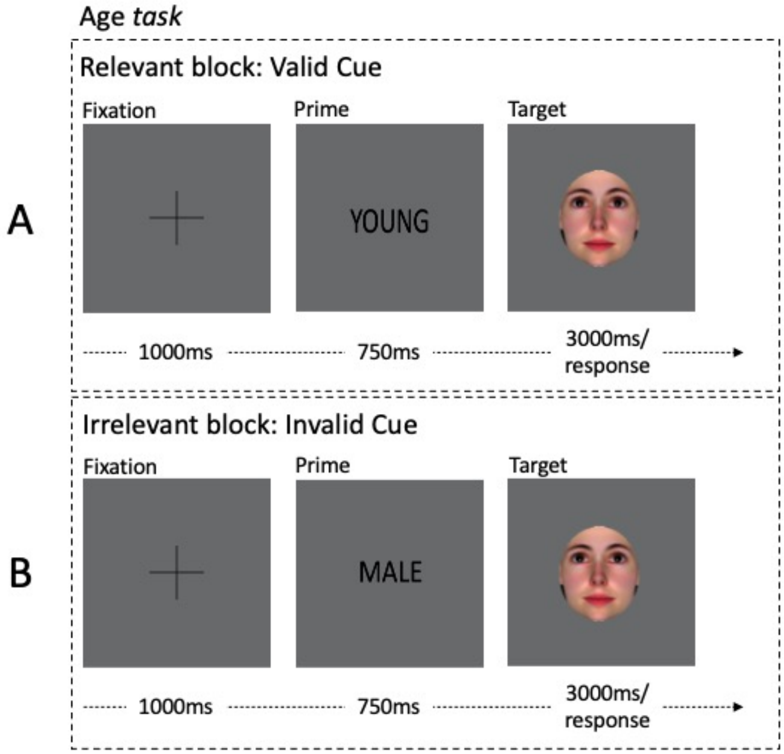


Figure 6.11: An example of each of the two block types for the Age *task*. (A) shows the Relevant block type with a Valid *cue type*: the prime relates to the *task*-relevant dimension of Age, and is an accurate cue to the Age of the target. (B) shows the Irrelevant block type with an Invalid *cue type:* the prime relates to the *task*-irrelevant dimension of Sex, and is an inaccurate cue to the Sex of the target.

6.4.3 Results

6.4.3.1 Control analysis

The control analysis was carried out in order to establish baseline performance for each *task* and *set* so as to inform the interpretation of the main experimental analyses. I collapsed IES scores on these factors across levels of distractor relevance and congruency and used a two-way between subjects ANOVA with *task* and *set*. This analysis revealed a significant main effect of *task* such that performance was better in the Age compared to the Sex *task* ($F(1,170) = 5.59$, $p = 0.01$, $h^2 = 0.032$, Sex = 766.15 (18.26), Age = 713.11 (16.33)).

These findings were qualified by a significant interaction between *task* and *set* (F(2,170) = 13.85, $p < 0.001$, h² = 0.14) (see figure 6.12). Follow-up analyses showed significant differences for the Facegen *set*, such that performance was significantly better in the Sex compared to the Age *task* (t(56) = 2.4, p = 0.01, d = 0.62; Sex M = 673.4 (27.9), Age M = 772.88 (30.7)), and for the MPI set such that performance was significantly better for the Age compared to the Sex *task* (t(46) = 4.73, $p < 0.001$, d = 1.26, Sex M = 842.43 (34.59), Age M = 647.38 (22.41)). There was no significant difference in performance between *tasks* for the Lifespan *set.*

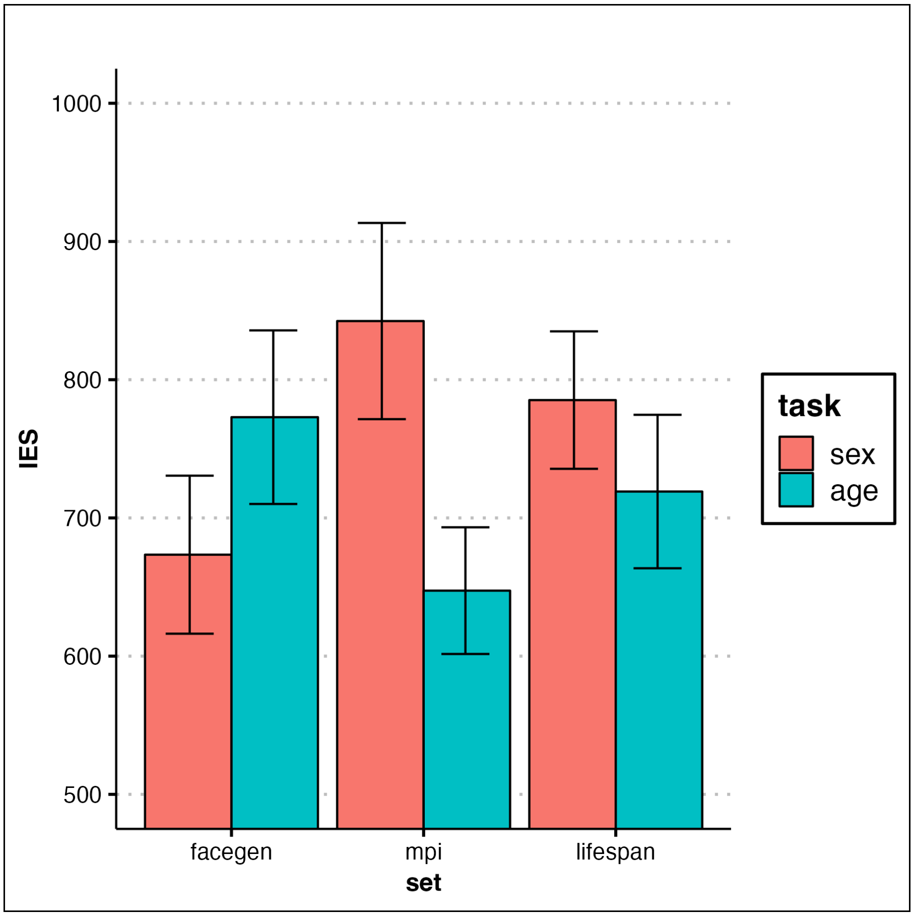


Figure 6.12: Control analyses for each *set* split by *task*.

6.4.3.2 Priming Effects

The presence of a priming effect was interpreted as a significant positive or negative difference score calculated by subtracting performance scores (IES) in Valid trials from Invalid trials. A positive difference score would be taken as evidence of a cueing benefit from valid relative to invalid cues, and a negative difference score would be taken as evidence of a performance cost for valid relative to invalid cues. Note that *cue types* either pertained to the *task*-Relevant *or the task*-Irrelevant dimension. That is, I separately measured the effects of verbal cues about the attended or the ignored dimension of the target face.

To assess the priming effect and its possible modulation by *relevance, task* and by stimulus *set*, difference scores were submitted to a three-way mixed ANOVA with *relevance* (Relevant or Irrelevant) as a within-subjects factor, and with *task*, and *set* as between-subjects factors. Prior to this analysis, participants whose difference score (averaged across conditions) fell outwith 2.5 SDs of the mean difference score were excluded from the priming analysis (n = 6).

The analysis showed a significant overall positive priming effect, interpreted as a y-intercept significantly greater than zero ($F(1,165) = 63.86$, $p < 0.001$, $h^2 = 0.16$; $M = 37.01$ (8.08)). There was a significant two-way interaction between *task* and *relevance* ($F(1,165) = 4.20$, $p = 0.04$, $h^2 = 0.01$) and a significant three-way interaction between *task, set* and *relevance* ($F(2,165) = 13.16$, $p < 0.001$, $h^2 = 0.07$). There was no main effect of *relevance* ($F(1,165) = 0.01$, $p = 0.9$) suggesting that valid semantic primes facilitated performance to an equivalent degree compared to invalid semantic primes regardless of the relevance of the prime to the task (see figure 6.13).

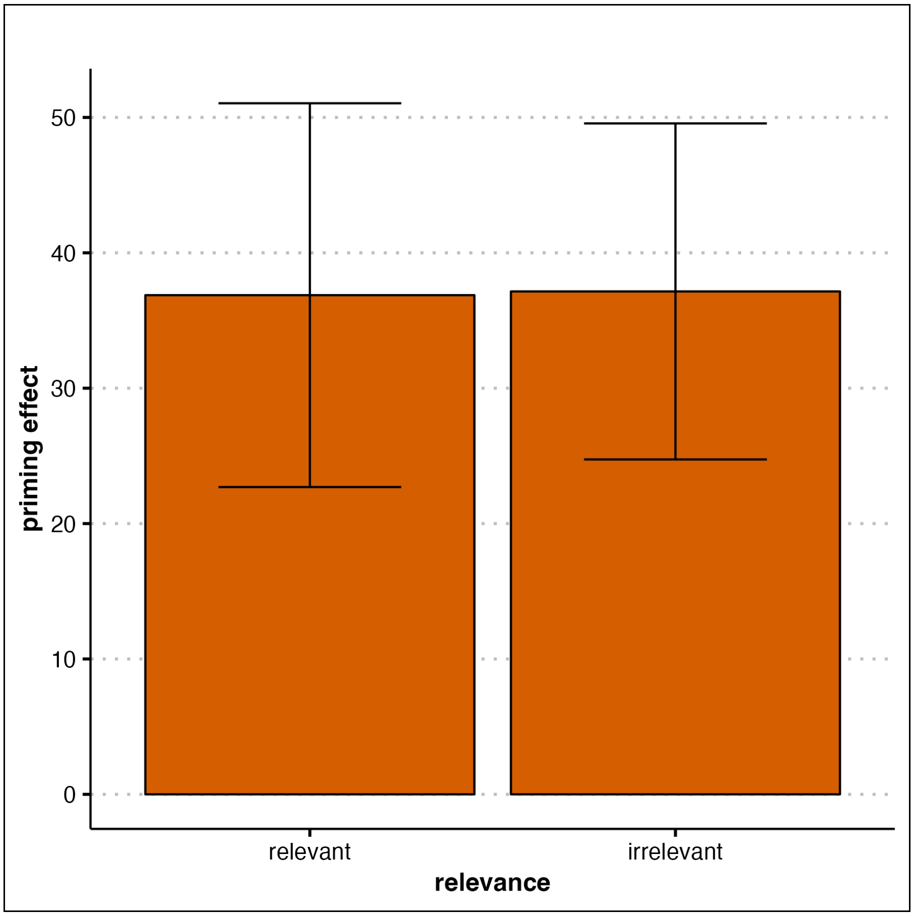


Figure 6.13: Priming effect for each level of *relevance*

Follow-up analysis of the three-way interaction involved carrying out three separate two-way mixed ANOVAs, one for each stimulus *set*. This analysis for the Facegen *set* showed a significant priming effect only (F(1,55) = 19.86, p < 0.001, h² = 0.17; M – 38.61 (1.60)) (see figure 6.14) but no other significant effects. This means that, for the Facegen *set*, the priming effect – the benefit of valid relative to invalid cues - was not mitigated by the task or the relevance of the cue that primed the target.

For the MPI *set*, there was a significant positive priming effect (F(1,55) = 18.04, p < 0.001, h² = 0.12; M = 29.85 (1.55)) and a significant interaction of *task* and *relevance* (F(1,55) = 6.86, p = 0.01, h² = 0.066). Follow-up t-tests for this interaction revealed a significant difference for the Sex *task* (t(27) = 2.45, p = 0.02, d = 0.73; Relevant M = 0.97 (3.36), Irrelevant M = 65.55 (2.97) but not for the Age *task*. That is, for the Sex task, a positive priming benefit was found for task-irrelevant Age cues but not for task-relevant Sex cues. In contrast, for the Age task, the priming benefits of Age and Sex cues did not differ significantly – although the difference trends such that Age cues were more beneficial to performance (i.e. a larger positive priming effect).

For the Lifespan *set,* there was a significant priming effect (F(1,55) = 18.04, $p < 0.001$, h^2 = 0.12; M = 29.85 (1.55)) and a significant interaction of *task* and *relevance* (F(1,55) = 6.86, p = 0.01, h^2 = 0.066). Follow-up t-tests were significant for both the Sex *task* (t(27) = 2.94, p = 0.006, d = 0.79, Relevant M = 65.61 (3.64), Irrelevant M = -6.08 (2.79)) and the Age *task* (t(26) = 2.78, p = 0.01, d = 0.86; Relevant M = 11.96 (2.78), Irrelevant M = 75.86 (2.74)). For the Sex task, a significantly higher positive priming effect was found when the cue was Relevant (Sex cues) compared to Irrelevant (Age cues). In contrast, for the Age task, a significantly higher priming effect was found when the cue was Irrelevant compared to when it is Relevant, i.e. for Sex cues relative to Age cues.

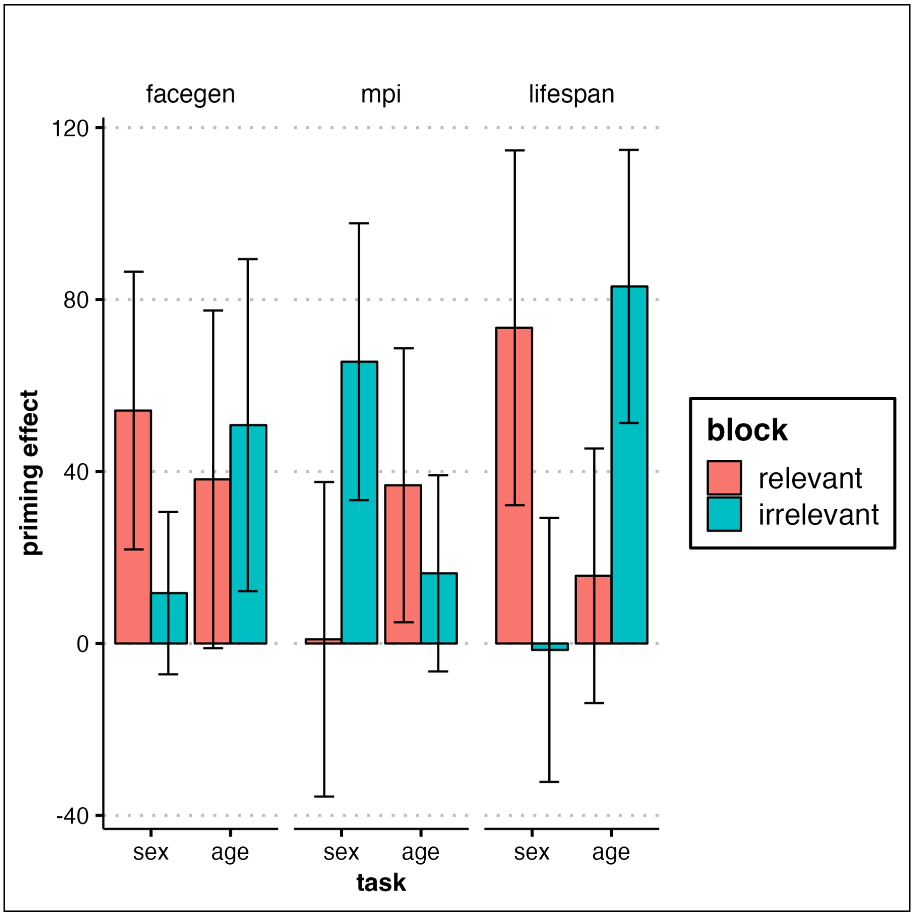


Figure 6.14: Priming effect for each *task* split by relevance for each *set.* Notice the difference between the MPI and Lifespan *sets* which characterise the three-way interaction.

Looking at the interaction between *relevance* and *task*, the follow-up independent t-tests were not significant for either the Sex *task* (t(82) = 0.73, p = 0.34) or the Age *task* (t(81) = 1.35, p = 0.18) (see figure 6.15). The interaction was characterised by an opposing pattern of difference between the levels of *relevance* for each *task* – so, as above, in general sex cues elicit higher priming effects than age cues, though this is complicated by the three-way interaction.

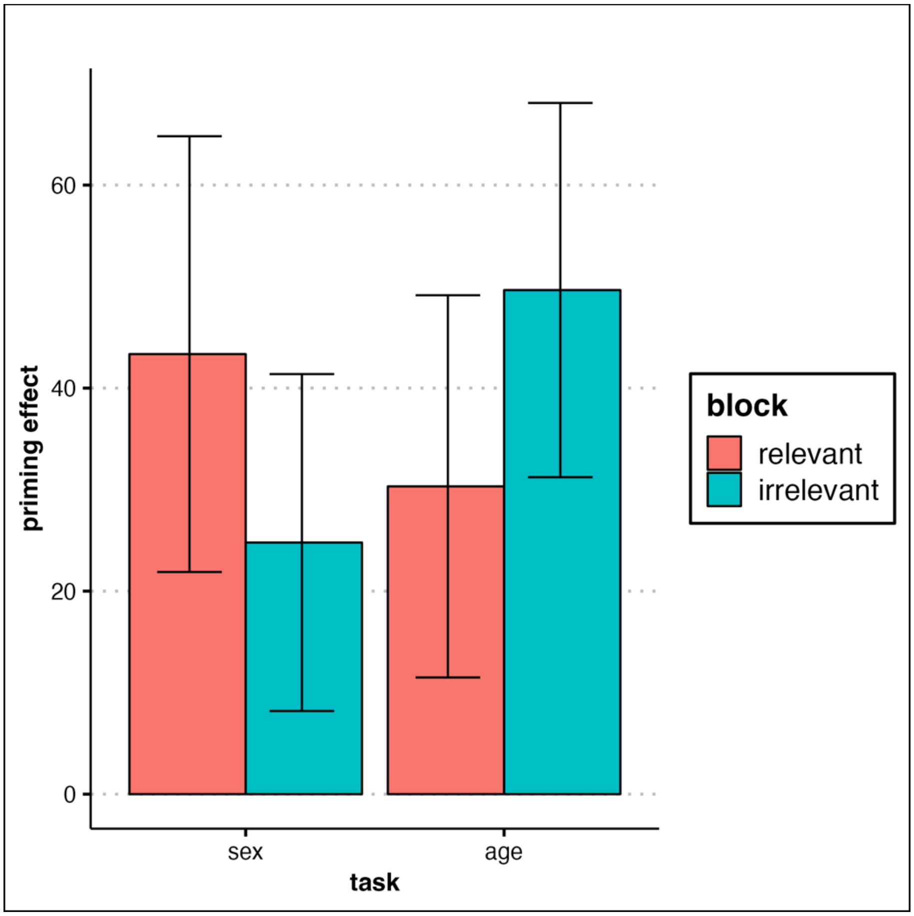


Figure 6.15: Priming effect for each *task* split by *relevance.* The interaction between the two factors is characterised by the reversal of the pattern of differential priming effects for each level of relevance between *tasks*.

6.4.3.3 Stroop Effects

Difference scores were calculated by subtracting the Compatible from the Incompatible configurations (see chapter 3). A positive difference score indicates better performance in the

Compatible compared to the Incompatible condition, and is interpreted as the presence of Stroop effects. Stroop effects and their possible modulation by prime *relevance, task* and stimulus *set* by submitting difference scores to a 2x2x3 three-way mixed ANOVA with *relevance, task* and *set* as factors. Prior to analysis, participants whose difference score (averaged across conditions) fell outwith 2.5 SDs of the mean difference score were excluded from the Stroop analysis (n = 9) The outcome of this analysis was a significant overall positive Stroop effect, interpreted as a y-intercept significantly different from zero ($F(1,162) = 119.87$, $p < 0.001$, $h^2 = 0.347$; M = 67.34 (7.83)). There was a significant main effect of *relevance* such that the Stroop effect was significantly larger in the Irrelevant compared to the Relevant blocks ($F(1,162) = 4.47$, $p = 0.04$, $h^2 = 0.008$; Relevant M = 58.78 (7.53), Irrelevant M = 75.9 (8.09)) (see figure 6.16). There was also a significant interaction between *task* and *set* ($F(2,162) = 10.83$, $p < 0.001$, $h^2 = 0.088$) (see figure 6.17).

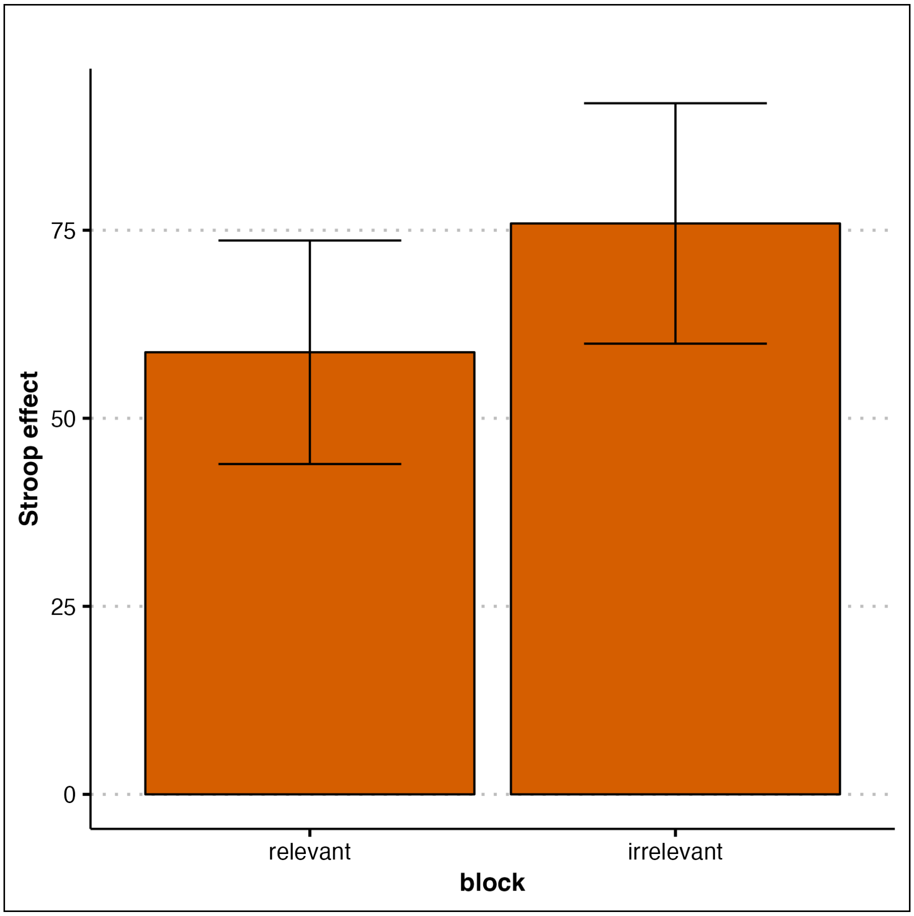


Figure 6.16: Stroop-like effect for each level of *relevance*

Follow-up analyses showed a significant effect of *task* for the Facegen *set* such that the Stroop effect was larger for the Age compared to the Sex *task* (t(97) = 5.42, p < 0.001, d = 1.05; Sex M = 44.73 (10.28), Age M = 134.85 (14.81)). For the MPI set the effect was significantly larger in the Sex compared to the Age *task* (t(69) = 2.97, p = 0.004, d = 0.6; Sex M = 91.64 (19.71), Age M = 47.48 (8.88)). There was no significant effect of *task* for the Lifespan *set* (t(94) = 1.54, p = 0.13).

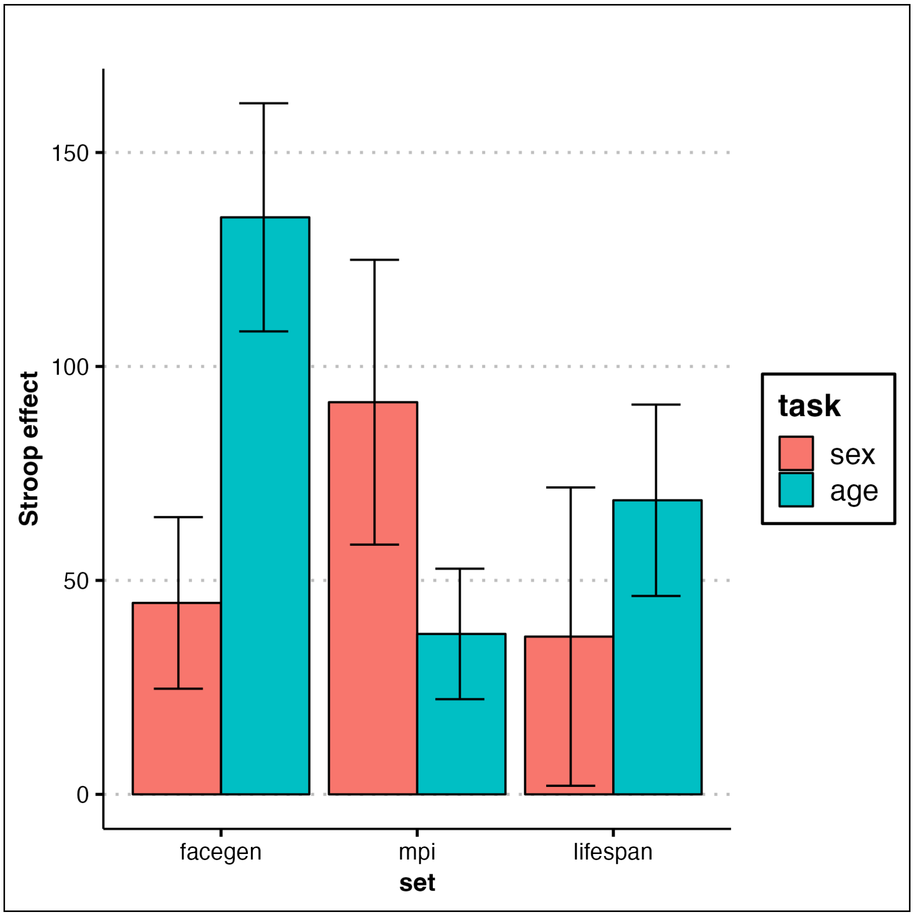


Figure 6.17: Stroop-like effects for each *set* split by *task*.

6.4.3.4 Cue-compatibility effect

The presence of a larger Stroop effect in the Irrelevant condition suggests that Compatible cues to target judgements were driving the effect. In order to explore this possibility. I looked only at valid, Irrelevant cues and created a difference score by subtracting scores on trials with Compatible cues (“MALE” cue to “Young” target and “FEMALE” cue to “Old” target in the Age *task*; “OLD” cue to “Male” target and “YOUNG” cue to “Female” target) from scores on trials with

Incompatible cues. A positive score would show that being primed with an Irrelevant but compatible cue led to better performance than being primed with an Irrelevant, Incompatible cue – an effect designated the cue-compatibility effect.

I submitted this difference score to a 2x3 ANOVA with *task* and *set* as factors. Prior to analysis, participants whose difference score fell outwith 2.5 SDs of the mean difference score were excluded from this analysis ($n = 2$), note that these exclusions were from the total number of participants, not on top of the 9 that were independently excluded from the Stroop analysis). This analysis showed a significant, overall, positive cue-compatibility effect – interpreted as a y-intercept significantly different from zero ($F(1, 170) = 42.97$, $p < 0.001$, $h^2 = 0.2$, $M = 83.62$ (13.25)). There was a significant main effect of *set* ($F(2,170) = 3.4$, $p = 0.03$, $h^2 = 0.038$, Facegen $M = 124.1$ (17.21), MPI $M = 84.36$ (29.89), Lifespan $M = 42.43$ (17.85) (see figure 6.18). FLSD (62.31) shows that there was a significant difference between the Facegen and Lifespan *sets* only.

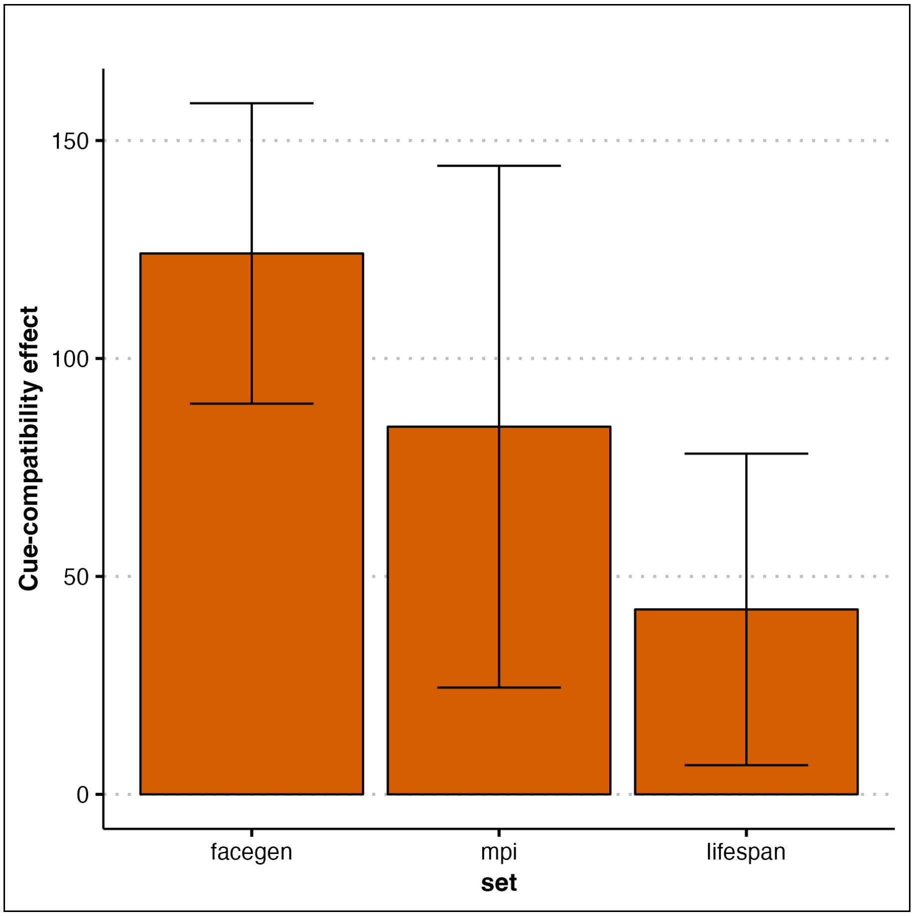


Figure 6.18: Cue-compatibility effect for each *set*.

6.4.4 Discussion

The current results show that a semantic prime to a general social category can affect performance in face categorisation. This is evidence of the top-down modulation of social categorisation and suggests that access to semantic information about an upcoming face makes it easier to categorise. There was no difference in the overall priming effect between task-relevant and task-irrelevant cues meaning that a valid cue to the age of an upcoming face facilitates a judgement about its sex as well as its age, and a cue to sex facilitates age as well as sex judgements. The results also suggest that the priming effect was largely driven by trials in which the primed dimensional value and the target dimensional value were compatible. Since *task* was a between-subjects factor, these results taken together rule out the possibility that priming occurs at the response selection level.

I also found significant Stroop effects which are comparable to those of the above reported face-priming study. This means that targets were processed more easily when their dimensional values were correlated. In the current case, I found that these effects were mitigated by the *relevance* of the cue – such that the effect was larger when the target was preceded by a task-irrelevant cue compared to a task-relevant one. Taking this further, I found that priming an irrelevant value that was compatible with the target value facilitated performance – suggesting that the activation of a semantic representation of a compatible value facilitates perceptual processing.

The interpretation of these effects is largely dependent upon a careful analysis of the similarities and differences between the three stimulus *sets* used and the baseline discriminability between the *task*s for each *set.* Such analysis suggests that the top-down modulation through semantic information interacts with the bottom-up modulation of stimulus features. I will briefly review the findings for the priming, Stroop and cue-compatibility effects, offer a holistic interpretation, and then explore the methodological issues and limitations of the experiment before suggesting directions for future research.

6.4.4.1 Priming findings

Valid primes were associated with better performance than Invalid primes *whether or not the prime was task-relevant*. This finding suggests that semantic information about an upcoming face can affect the processing of that face, even if the semantic information is not relevant to the processing taking place. Whilst the three-way interaction between *task-relevance*, *task* and *set* indicates that the extent of this priming varied between factors, the effect itself was robust, and there was very little main effect of *relevance*.

This means that predictions generated by access to semantic information can affect perceptual decision-making, and that this occurs even when that information is ostensibly irrelevant to the target about which the decision is made. This latter fact can be taken as further evidence of perceptual integrality between sex and age information, since information about one affects the processing of the other. If sex and age were separable dimensions, predictions about one would not affect processing of the other.

The extent of this interaction may vary as a function of the extent to which the properties characterising the levels of each dimension overlap. That is, a prediction based on a prime to the Irrelevant dimension may permit some degree of "pre-processing" of that dimension – some of which happens to be relevant to the task-relevant judgement due to overlapping or mutual sensory or semantic properties.

An example which might account for this phenomenon in visual terms involves the adoption of a top-down processing strategy tuned to process the primed dimensional value, "Old", for example. Even when that value is irrelevant, its facilitated processing permits the processing of the relevant value to be carried out more efficiently – without the irrelevant value requiring the level of visual attention that would ordinarily be necessary in the absence of a prime. If sex and age are integral then one cannot be processed without some processing of the other, and thus if the processing of one is facilitated, the processing of the other should be facilitated as well. Alternatively, performance may be inhibited if one is primed to attend to a certain pattern of features but then the subsequent visual display contains a different visual pattern – consistent with a different value. One then has to disengage the primed top-down visual processing strategy and process the image from scratch – which is presumably less efficient than processing from scratch to begin with. This would explain relatively poor performance with invalid compared to valid primes.

Another possible explanation relies upon underlying semantic (that is, conceptual) similarities between values across dimensions – such that the activation of one as a function of the prime activates the related dimension through a mechanism like spreading activation (Collins & Loftus, 1975). This interpretation also works both ways, explaining facilitation in the case of valid primes and inhibition in the case of invalid primes – since in the former case there is less activation necessary to trigger categorisation, and in the latter case activation has to decrease for the invalid and increase for the valid value. Semantic representations of dimensional values pertaining to sex and age may themselves are multidimensional, and may be characterised by their own patterns of mutual integrality with other representations.

These interpretations, which posit integrality at the level of sensory and semantic representation respectively and which suggest that the properties of sex and age overlap in such a way that information about one affects the processing of the other, are somewhat complicated by the fact that the relationship appears to be asymmetrical. This asymmetry is highlighted by the interaction between *task* and *task-relevance*. Primes to Sex interfere (positively and/or negatively) more than primes to Age, regardless of the task at hand. This, however, is complicated further by the fact that this asymmetry is not uniform and differs between *sets*.

That being said, there are distinct patterns of trends between *tasks* and levels of *task-relevance* within and between *sets*. For example, it is always the case in the current results that the extent of the priming effect is more for one *task* when the prime is Relevant, and more for the other *task* when the prime is Irrelevant – even if the extent of this difference and the direction of that difference changes between *sets*. The consistency of this finding for all three *sets* is evidence of asymmetrical integrality in the processing of Sex and Age information from faces, however a more specific conclusion is difficult to draw given the inconsistencies between *sets*.

The priming findings overall evade interpretation according to the SoPA, as they do not map consistently onto the patterns of baseline task-difficulty. For the Facegen and MPI *sets*, semantic information pertaining to the least discriminable (i.e. more difficult) *task* had the biggest impact on performance. For the Lifespan *set* it was the opposite, with semantic information to the least discriminable dimension having a bigger impact. For the Lifespan *set*, the positive priming effect is largest for Relevant cues to Sex and Irrelevant cues to Age (i.e. cues to Sex). In fact, cues to Age for the Lifespan *set* had no reliable impact on performance.

The opposite is true for the MPI *set*, which is somewhat surprising given that both the MPI and the Lifespan *sets* are photographic, and that they share a pattern of task-difficulty (though the difference in difficulty between *tasks* is significant for MPI and not Lifespan). One might therefore expect that the difference in apparent asymmetrical integrality would be one of degree, not a complete reversal.

This fact is difficult to reconcile into a coherent, general interpretation, and serves as a further warning not only against interpreting results based on judgments about a single set of stimuli, but about interpreting results with multiple sets of stimuli in relation to only one experiment. The pattern of results observed in the analysis of the priming effect were mirrored in the Stroop analysis, such that the Facegen and Lifespan *sets* shared a pattern of results which was opposed by the MPI *set.* Differences between *sets* prevent us from making strong claims as to the underlying organisation of the face perception system, instead perhaps pointing more strongly to the influence and importance of stimulus features in the outcome of integrality experiments.

The priming effects findings show that top-down semantic information can affect social categorisation from visual face perception, that the *relevance* of the semantic information to the social categorisation itself does not affect performance per se, only as a function of the specific nature of the faces about which the categorisation judgement is to be carried out. The findings suggest that semantic information pertaining to the dimension of sex can affect performance to a greater degree than semantic information pertaining to the dimension of age – in keeping with the hypothesis that sex perception operates from the top-down more so than does age perception. Differences between *sets* prevent us from making a strong claim to this effect based on the current data alone, however previous experiments within the current volume bolster the interpretation.

6.4.4.2 Stroop and cue-compatibility findings

The Stroop analysis gives us a measure of the effect of task-irrelevant but perceptually compatible information on performance. The Stroop effect was larger in the Age *task* than the Sex *task* for both the Facegen and the Lifespan *sets* (although the difference was not significant for Lifespan) whereas it was larger in the Sex compared to the Age *task* for the MPI *set.* This means that, for example, covert and embedded cues to Sex have a bigger impact on performance in the MPI *set* than do such cues to Age.

I found that the Stroop effect was larger in the Irrelevant compared to the Relevant *blocks* – whereas I did not find that the priming effect varied as a function of *relevance*. What this effect tells us is that a prime to an irrelevant value can actually have a greater impact on performance than a cue to the relevant value. Following this up with the cue-compatibility analysis, I found that an irrelevant but compatible cue facilitated performance relative to an irrelevant and incompatible cue. This varied as a function of *set* but not of *task.* The effect was significantly higher for the artificial Facegen *set*. This may be because the stimuli are relatively unfamiliar and are unlikely to be as ecologically valid as the photographic *sets* – perhaps increasing the likelihood of the recruitment of semantic (top-down) information in the rendering of a judgment.

Interestingly, the cue-compatibility effect was not modulated by *task*, but the overall Stroop findings were. The cue-compatibility effect is more of a measure of the impact of target-compatible semantic information and measures the impact of overt-unembedded information, whilst the Stroop effect is more of a measure of target-compatible visual information and measures the impact of covert-embedded information. The discrepancy between the two effects in terms of their modulation by *task* may point to differences in the way that these two types of information interact with social categorisation judgements.

6.4.4.3 Interpretation

The current results provide further evidence for asymmetrical integrality between the dimensions of sex and age. Performance in categorising faces according to these dimensions is differentially affected by semantic information and the generation of top-down predictions (expectations). Taken as a whole, the experiment suggests that cues to sex affect processing more than cues to age, however differences between stimulus *sets* suggest a role of the physical properties of the stimuli in modulating the effect. The impact of semantic information on face processing may vary as a function of the nature of the available visual information.

I can tentatively argue that the activation of a representation of sex provides access to more complex, configural templates upon which to base a judgement – facilitating processing regardless of *task*. When an age representation was primed, only simpler, featural templates become active. These fail to facilitate sex judgements to the same extent because the “deeper”, more complex, configural processing still needs to take place, and they fail to facilitate age judgements to the same extent because featural templates are more “deformable” and less static (see Yuille, 1991). That is, age-related templates are less universal – because people age differently – and so matching bottom-up information to an active age-template may require more processing than matching such information to a sex-template – a template which constrains that bottom-up processing by acting as a structural referent for it.

According to such an account, semantic cues to sex activate sex representations which constrain the processing of age. Representations of sex are also more consistent across faces than representations of age, and bottom-up sensory information about the target face is therefore more easily matched to the sex-template than the age-template. This can also be thought of in terms of parallel-contingent processing. Sex and age judgements are contingent upon the activated representations of sex, but they are less so upon active representations of age.

If this were the case, however, we would expect the effect to be more consistent across *sets*, but they are not. It may be that the semantic information tunes or prepares the visual system through recurrent feedback to receive information consistent with the cue (see Gilbert & Li, 2013). According to Ramalingam, McManus, Li and Gilbert (2013), early visual cortex areas including V1 change their computations according to the perceptual task at hand. V1 neurons have been described as “dynamic processors” (Kamiyama, Fujita & Kashimori, 2016), changing their function according to the relevant behavioural context.

Henderson, Tarr and Wehbe (2023) recently found evidence to support the idea that representations of visual and semantic information can overlap in cortex, with the same neural populations being sensitive to low- and high-level information. They argued that low-level feature selectivity contributes to the computation of semantic, category level information in the brain. They showed associations between responses for low-level (Gabor) features and high-level semantic categories, and showed that voxels in category-selective regions exhibited biases towards selecting feature- and spatial-level information consistent with the category to which they were selective.

It may be, then, that the provision of a semantic cue allows early visual processes to be tuned to the reception of a visual signal consistent with the semantic signal. That the activation of a high-level semantic category facilitates the processing of low-level visual featural and spatial selectivity because there is an overlap between the neural populations associated with each process. This would explain the finding that task-irrelevant but compatible information produces a positive priming effect and the finding that the word priming effect varied between *sets*. Although the current results can be interpreted in such terms, it may be wiser to interpret them as evidence of the sensitivity of the face processing system to differences in stimulus properties and the relative discriminability of facial dimensions.

Methodologically, then, the current study further highlights the fact that the set of stimuli used is not arbitrary, and that different patterns of results can emerge from different stimulus sets. The results also point to the possibility that semantic and sensory, or overt-unembedded and covert-embedded information interact with face perception in different ways, raising questions about potential interactions between these types of information.

6.4.4.4 Limitations and future directions

Any future replication of the current experiment should include a longer training phase, which would have allowed the participant to learn that the cue was typically valid. There were only four practice trials in which the cue was always valid; if this practice block were longer, it would have made it more clear to the participant to expect valid cues. As it stands, it may be that the participant only came to rely on the cue to facilitate their judgement at some later point within the experiment. The location of this point in the experiment is likely to have varied between participants. Because of these likely individual differences and because the blocks were shown in a random order, this issue might constitute a source of noise rather than a confound.

As to the priming analysis, the design of the current experiment means that I can only interpret the effects of Valid primes and Invalid primes with respect to one another. That is, I did not include a control task to treat as a baseline, an example of which might have been to prime some trials with non-social categories which were neither Valid or Invalid descriptors of the target faces – although generating examples of these which were semantically and affectively neutral in nature might be difficult. If I had done that, I would be able to determine whether the differences in performance noted between Valid and Invalid primes were the result of facilitation or inhibition or both facilitation and inhibition. As it stands, I can only say that performance was better when primes were Valid compared to when primes were Invalid. Inclusion of a control condition might even reveal that there are asymmetric patterns of facilitation and inhibition.

The current experiment suggests possible differences between the ways that semantic and visual or overt-unembedded and covert-embedded information affects face perception. A way to probe this effect more deeply would be to replicate the current design but use faces as primes and words as targets rather than the other way around. This would permit us to see whether faces – and in particular, the irrelevant information about them – would affect categorisation judgements about words in the opposite way as words have been shown to affect categorisation judgements about faces.

6.4.4.5 Conclusion

The current results show clearly that semantic information can affect perceptual decision making about faces – even, and in some cases especially, when the semantic information is not directly relevant to the task at hand. The results also suggest that there may be asymmetries in the way that sex and age are processed visually and semantically, and that the nature of these asymmetries might not be the same between domains (the sensory and the semantic domain) – though these interpretations are speculative and are not lent sufficient weight by the current data to make a strong claim. The experiment also provides a warning against interpreting social vision experiments using only one set of stimuli, since it appears that the processing of different *sets* can lead to vastly different results, even when they are both photographic.

6.4 General Priming Discussion

Taken together, these experiments show that the social categorisation of human faces can be affected by bottom-up and top-down modulation, that socially relevant information can impinge upon social categorisation in an automatic as well as in a volitional manner, and that sex and age are asymmetrically integral dimensions. There are several similarities and differences between the experiments which tend towards a common interpretation, but which simultaneously warn against overgeneralising from effects measured in relation to any given set of stimuli. I will discuss the findings of both experiments together and offer an interpretation, I will explore how future research in this area might address some of the limitations of the current experiments.

6.4.1 Baseline Discriminability and the Stroop Effect

The baseline discriminability between *tasks* for each *set* were similar across experiments – with the Sex *task* being more discriminable for the Facegen *set* and the Age *task* being more discriminable for the MPI and Lifespan *sets*. The profile of Stroop effects also followed the same pattern between experiments, with a non-significant difference between *tasks* in the Lifespan *set* and significant differences such that the effect was larger for the Sex *task* for the MPI *set* and vice versa for the Facegen *set*.

These findings suggest a connection between Stroop effects and baseline task discriminability. Where task discriminability is significantly different, larger Stroop effects emerge for the least discriminable task. Small differences in *task* discriminability are associated with small disparities in the Stroop effect between *tasks* (i.e. for the Lifespan *set*). Following the logic of the SoPA, one interpretation would be that there is more scope for the more discriminable dimension to interfere with the less discriminable dimension. It may be that it is possible to make the easier (more discriminable) judgement earlier, and before the interference from the irrelevant dimension manifests.

6.4.2 Priming Effects

The interaction between *relevance* and *task* was significant for both experiments – although this was modulated by an interaction between these variables and *set* in the word priming experiment, whilst this interaction was marginal in the face priming experiment. These interactions between *relevance* and *task* suggest that sex information has more of an impact over performance than does age information regardless of the *task*. This may suggest that sex information is more salient than age information, and that processing of sex information takes priority over the processing of age information regardless of the *task*. Put another way, sex information modulates attention in a bottom-up or top-down manner to a larger extent than does age information.

This is consistent with the idea that sex perception is a more complex, or “deeper” process than age perception, that sex acts as a structural referent for age, and that age perception is contingent upon sex perception. The current results speak to this hypothesis by showing that information about the sex of an upcoming target has a bigger influence over categorisation performance than does information about its Age. The deeper level of processing interferes with the "shallower" level of processing more so than vice versa.

The major difference between the two experiments relates to the finding that Congruent primes facilitated performance in the word priming experiment and inhibit performance in the face priming experiment. If any activation of the representation affected subsequent categorisation performance, we should expect similar results in both experiments. The different priming results between the two experiments may have to do with the level at which competition between processes takes place.

Although we can infer from the results that the information (sensory or semantic) from the prime is extracted, and – theoretically – we can suppose that this extraction of information activates representations of the dimensional values. When it comes to selecting between these representations, semantic information may facilitate the process whilst visual information may inhibit it. There are also some key differences between the types of primes used across the experiments which may have led to the difference in priming effects.

In the word priming experiment, the primes usually referred accurately to a dimensional value of the target, and they also contained information about only one dimension. The primes in the face priming experiment referred accurately to dimensional values of the target as often as they didn't, and also contained information about both dimensions of interest. The semantic primes are also abstracted and impersonal, referring to a general semantic category. In the face-priming experiments, the information is embedded within an identity – it is personal, and must be abstracted from the individual identity in which it is embedded. There is more information in the face primes than just the sex and age information measured in the analysis. These facts are unavoidable aspects of the nature of the stimuli used as primes, but there are ways that the two types of primes could be rendered more similar in future experiments.

If the semantic primes were "unreliable", that is, did not prime an upcoming aspect of the target more often than not, or if the face primes were "reliable" – typically displaying information that was congruent with the upcoming target - I may have found more similar results between the two experiments. In order to test the difference in dimensionality, it would be necessary to use two words (for example “Old Woman”) in the word priming experiment, since I could not reduce the dimensionality of a face to showing only sex or age relevant information.

6.4.3 The Effects of Stimulus Set

Turning now to the interaction between *relevance*, *task* and *set,* we see that the effect is greater for the word priming than the face priming experiment. Comparing the two effects more carefully (figures 6.8 and 6.14), we can see that the interaction in the word priming experiment was characterised by greater differences in extent and direction of the priming effect between levels of *relevance* and between *tasks* for each *set.* Specifically, the profile of effects seems more similar between the Facegen and Lifespan *sets*, whilst the MPI *set* appears to display a near opposite pattern of results. In the face priming experiment, the differences seem largely to be differences in extent, with only minor differences in direction – for example between the levels of *relevance* of the Age *task* between the Facegen and MPI *sets* (figure 6.14). The profile of priming effects in the face priming experiment seems more heterogenous across *sets*. This modulation of the priming effects by *set* is in contrast to the Stroop effects – which were similar across both experiments and which can be interpreted in relation to the baseline task-difficulty.

6.4.4 Limitations

Apart from the limitations discussed above, it is worth re-stating that these two experiments are not perfect counterparts. They require different patterns of behaviour from the participants. They are two studies aimed at addressing complementary questions in different ways, and cannot be seen as a dyad of experiments which perfectly converge on the core issue of attentional modulation. Future projects looking to explore bottom-up and top-down modulation of the same stimuli may aim at homogenising the two experiments to make them more directly comparable.

Another difference that may have caused a discrepancy in the way that the primes affected the processing of the targets relates to the stimulus onset asynchrony (SOA) between the prime and the target. In the word priming experiment, there was no delay between the offset of the prime and the onset of the target, but in the face priming experiment there was a 500ms delay. Both facilitation and inhibition might be expected with long SOAs, which 500ms could be considered to be, whereas only facilitation due to spreading activation is expected with shorter SOAs (Marí-Beffa, Hayes, Machado & Hindle, 2005). Equalising these SOAs between experiments would be advisable in future experiments.

6.5 Conclusions

Social categorisation of faces according to sex and age can be subject to both bottom-up and top-down modulation of attention. That is, automatic and volitional control affects social categorisation performance. These different types of modulation affect performance differently, with valid semantic (top-down) information facilitating judgements relative to invalid semantic information, and valid visual information (bottom-up) typically inhibiting performance relative to invalid visual information.

Taken together, the experiments suggest that sex information has a greater impact over performance than does age information, regardless of whether the task at hand involves a sex or an age judgement. This finding is consistent with the structural-reference hypothesis (Ganel et al., 2004) and the parallel-contingent hypothesis (Johnstone & Downing, 2017), such that sex acts as a structural referent for age and age processing is contingent upon sex processing.

These two experiments further emphasise the danger of using a single stimulus *set* to explore issues of perceptual integrality, showing again that the results can vary quite widely between them. Carrying out two studies, each with multiple stimulus *sets* which remained constant across experiments, has allowed for a consistent picture to emerge which – although noisy – offers a reliable interpretation of the interaction between the two studied visual dimensions.

CHAPTER 7 Task-switching as a measure of perceptual integrality

7.1 Abstract

In this chapter I investigated the integrality of sex and age using a task-switching paradigm in which participants were expected to switch between sex and age processing at regular intervals. This paradigm tells us about the relative costs and processing loads associated with conducting each task and offers insight into how sex and age interfere with one another at perceptual levels as well as at the response levels. I measured general and specific switching costs, which have been proposed to index the processing decrement associated with keeping an irrelevant task-set in working memory, and the effects of competition associated with the irrelevant task's response schedule respectively. I found asymmetric costs which were modulated by task-difficulty for the general switch cost, and symmetrical specific switch costs. This suggests that the effect of keeping an irrelevant task-set in working memory is modulated by how difficult the associated task is to perform, and that response level competition between sex and age occurs, but any asymmetry in their interference has been resolved by the stage of response selection. I found a robust Stroop effect for each stimulus set.

Acknowledgements

I would like to thank Dr Paloma Mari-Beffa for her help in understanding some of the theory behind task-switching, for her advice in interpreting the results and for suggesting the post-hoc analysis – the results of which turned out to be instrumental in the overall interpretation of the findings.

7.2 Introduction

The complex and dynamic nature of our environments necessitates the capacity to behave adaptively – we need to be able to flexibly engage and disengage our attention. Our ability to do this is often described in terms of cognitive control or cognitive flexibility (see Koch et al., 2018 for a review), and these faculties are often investigated through the limits of our ability to switch between cognitive tasks – known generally as *task-switching* (see Monsell, 2003 and Kiesel et al. 2010 for reviews). Intuitively, if a person is reading a paper whilst exchanging e-mails, they may find that switching from reading, to e-mailing to reading again may cause them to lose momentum in their reading – being unable to return to it from an e-mail with the same level of attention that they left it with. In terms of the theory of task-switching, this is because each of the two tasks – reading and sending a message – have associated *task-sets*, patterns of perceptual, cognitive and motoric demands associated with the specific task which are interrupted during the switch.

In the task-switching literature, a *switching cost* is thought to be a measure of the dissipation of the task-set associated with the previous task, the preparation of the new task-set, and the 'residual component' - or the interference of the old task-set with the new one (Meiran, Chorev & Sapir, 2000). The old task-set must be disengaged, the new one engaged, and interference between them must be resolved – these are components of the switching cost. In terms of the example above, one would have to disengage from the paper and switch to sending an e-mail. The switch might involve stopping reading and starting writing. The semantic content of the two tasks might interfere with one another – for example one might set out to respond in the e-mail to something that was read in the paper.

Task-switching in the lab typically involves rapid switching between two or more forced choice reaction-time tasks, and is usually associated with the decrement in performance known as a switching cost (e.g. Biederman, 1972; Jersild, 1927; Wylie & Allport, 2000). Despite its long history, research into the capacity to switch between tasks is relatively poorly integrated (Koch et al., 2018). There seems to be a wide variety of different techniques employed and theoretical approaches taken towards understanding the phenomenon of switching costs, and investigations have been aimed at understanding the mechanisms underpinning switching costs, but also as tools to shed light on several other perceptual and cognitive problems. This has led to a lack of terminological and theoretical consistency within the literature. To give an example at the highest level, the ability to change behavioural goals has been referred to as task-switching, set shifting and attention switching (Ravizza & Carter, 2008).

The task-switching experiment reported here involves switching between two speeded classification tasks, one in which participants must classify faces by their sex, the other by their age. In previous chapters, we have seen an asymmetrical relationship between performance on these dimensions. Sex information seems to have a larger impact upon age judgements than vice versa. However, to this point I have not directly compared tasks - *task* has so far been a between-subjects variable. To discuss what the task-switching paradigm can tell us about the processing of these dimensions, I will start with a brief discussion of the paradigm and the theory behind task-switching and switching costs, before exploring the work that has already focused on these dimensions.

7.2.1 The task-switching paradigm

Ravizza and Carter (2008) set out to try and define some of the terms in the task-switching literature, to delineate between different aspects of the process of switching tasks (setting up the new task-set, disengaging the old one and resolving interference), and to determine whether these different aspects were underpinned by different neurological mechanisms. They make a distinction between *set-shifting* and *task-switching.* Set-shifting has been defined in terms of changes in “attentional set”, with *set* being defined as the property of the stimulus that is relevant on a given trial (Rushworth, Passingham & Nobre, 2005). In a visual set-shifting experiment, one must attend to the set of visual properties that are relevant to one dimension of the stimulus, and then shift to the set relevant to the other. Attentional shifting is switching your attention to a different object or aspect of an object.

Task may refer to a change in the goal state associated with the outcome of the operation. Task-switching can therefore be defined in terms of changes in “intentional set” (Rushworth et al., 2005). Being instructed to attend to one dimension of the stimulus constitutes being instructed to conduct one *task* – or engage one set of intentions, shifting to another intentional set is akin to switching to another *task*. This can be more broadly defined to include knowledge and memory for the responses and stimulus-response mappings for both the relevant and the irrelevant task – thereby implicating working memory (Rogers & Monsell, 1995). The intentional and attentional *sets* can be defined in terms of the *task-set* – the efficient organisation of the perceptual and cognitive system around the performance of the *task* (Yeung & Monsell, 2003).

Ravizza and Carter (2008) equate the definitions of *set* and *task* with “perceptual switching” and “rule switching” respectively, pointing out that most task-switching experiments conflate these two types of switch without accounting for the difference. In our example of switching between reading and sending an e-mail, that process would require set-shifting from the paper on the desk to the e-mail on the laptop, it would also require switching between the task of holding up the paper and paying attention to the written words written on the paper to operating the keyboard and conveying what one wanted to write. One would have to switch their perceptual focus and the cognitive and motoric operations – or intentional and attentional sets respectively - in order to effectively do each task. Taken together, we would refer to this as the *task-set*.

Ravizza and Carter (2008) go on to demonstrate that perceptual shifting or set shifting is associated with greater engagement of the parietal cortex. Rule switching or task-switching is associated with increased engagement of the dorsolateral prefrontal cortex. From this evidence we can conclude that there exists a neurological dissociation between perceptual and response level operations in task-directed behaviour. From the experiments reported above (chapters 4, 5 & 6) it appears as though asymmetric interference between sex and age arises at the perceptual level but not the response level – although the two processing tasks have not yet been put into direct competition at the response level as they are in the current experiment.

Task-switching experiments typically require participants to learn two tasks and sets of instructions and responses (two intentional sets). Both tasks are usually undertaken with the same stimuli, although the precise features of those stimuli to which the participant must attend may vary between tasks (attentional sets). Between trials, the task sometimes changes and sometimes stays the same. It has consistently been noted in task-switching experiments that performance is worse on *task-switch* (or *switch*) compared to *task-repeat* (or *nonswitch*) trials: that is, trials in which the task has just changed compared to those in which the task is the same as on the previous trial.

Interestingly, switching between two different tasks often comes with asymmetrical switch costs, such that it is more difficult to switch from one to another than vice versa (e.g. Allport, Styles & Heisch, 1994; Ellefson, Shapiro & Chater, 2006; Wu et al., 2015). This makes the task-switching paradigm a uniquely useful if underutilised tool for investigating the perceptual integrality of complex multidimensional visual objects, including faces, where there is reason to expect asymmetrical effects. Debate persists as to what underpins asymmetric switching costs, although it is generally agreed that they are associated with *task dominance* (Gilbert & Shallice, 2002; Hautekiet, Verschooren, Langerock & Vergauwe, 2023).

This is the idea that one task is more dominant (better learned, more automatic and easier) (Gilbert & Shallice, 2002) than another. Counter-intuitively, it has often been shown that switching to the more dominant (easier) task is more difficult than switching to the less dominant (harder) task (Allport, Styles & Hsieh, 1994; Ellefson, Shapiro & Chater, 2006, but see Monsell, Yeung & Azuma, 2000). A number of potential mechanisms have been forwarded to account for such asymmetrical switching costs. These accounts, which can be sorted in terms of bottom-up and top-down control, are reviewed next.

7.2.2 Asymmetric switching costs

Allport et al. (1994, 2000) argued that the switch cost reflects the time required to resolve interference from the persistent activation of a previously active *task-set* – an idea known as *task-set inertia*, by which the effect of one *task-set* being active carries over and interferes with a new task-set once active – a phenomenon also known as the residual cost (Meiran et al., 2000). According to this account, when *task-set* A is much stronger than *task-set* B, it is necessary to actively inhibit A when performing B. When switching from B to A, this active inhibition persists and must be overcome – resulting in switching costs. Because the *task-set* for task B is weaker, it doesn't need to be inhibited, and inhibitory processing therefore does not carry over into performance on the other task.

Meuter and Allport (1999) conducted a task in which bilinguals switched unpredictably between naming digits in their primary (dominant) and secondary (nondominant) languages. Naming was faster in the dominant language in the non-switch trials compared to the nondominant language but slower in the switch trials – another example of it being easier to switch to the harder task. Despite the evidence for this counter-intuitive finding, it has been argued that it reflects a processing strategy adopted only when the difference in difficulty reaches a certain threshold at which it becomes necessary to engage in active inhibition (Monsell, Yeung and Azuma, 2000).

Yeung and Monsell (2003), conducted a standard Stroop task (Stroop, 1935) in which they made the more dominant task (word-naming) more difficult – but, crucially, not to the extent that it was more difficult than the nondominant colour-naming task. Under ordinary conditions they replicated the findings of Allport et al., (1994, 2000), that switching to the dominant task came with larger switching-costs. When they made the word-naming task more difficult, the pattern of switching costs was reversed – revealing that the surprising finding that the more difficult task is harder to switch to is not always the case.

Yeung and Monsell (2003) developed a computational model by which it was assumed that top-down control is also implicated in task-switching, ensuring that the relevant task-set is maintained whilst relevant and inhibited once irrelevant. More top-down control is needed to activate and maintain the nondominant task because it has a lower level of baseline activation. This model supports the idea that task-switching costs are often characterised by *task set inertia.* However, this is mitigated by the strength of top-down inputs in the nondominant task. The more top-down control, the less *task-set inertia* – that is, the easier it is to disengage the *task-set.* The more dominant the task, the less top-down control is required to initiate and maintain the task in working memory (see figure 7.1).

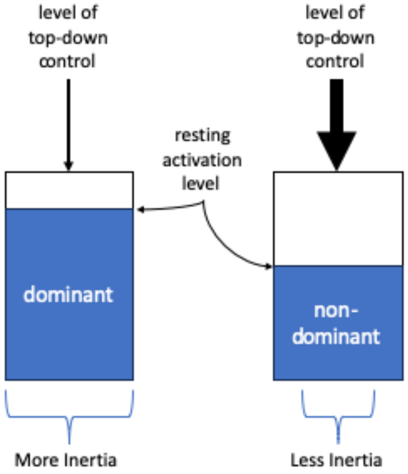


Figure 7.1: A diagram of Yeung and Monsell's (2003) computational model. The more dominant task has a higher level of resting activation, and therefore requires less top-down control to maintain it. The relative lack of top-down control over this task results in more task-set inertia once the task is complete and needs to be switched. The opposite is true of the non-dominant task; a lower resting activation level requires more top-down control, which results in less inertia.

Taking the above ideas together, we can say that there is a dissociation between task-switching at the perceptual and response levels (Ravizza & Carter, 2008), and that the "dominance" of the tasks at hand is implicated in what makes switching more or less costly at these levels. Switching costs are characterised by setting-up a task-set, the inhibition or dissipation of the old task-set, and interference between the two tasks (Meiran et al., 2000). These aspects of the process involve some combination of the resting rate of activation of a task-set (intentional and attentional sets) and the top-down control brought to bear on those task-sets (Allport et al., 1994; Yeung & Monsell, 2003). Switching to a new task may involve the active inhibition of the old task, but perhaps only when the difference in dominance between two tasks is sufficiently large.

What should we expect to find, then, given the design of the current experiment? My previous experiments (chapters 4 and 6) have thus far suggested that there is an asymmetric relationship between sex and age processing, but they don't necessarily tell us which we should expect to be more dominant. In the photographic *sets*, age processing has been found to be easier, but tends to suffer from more negative interference from sex information than does sex processing from age information. It was also argued, based on the results of the priming experiments (chapter 6), that sex processing takes priority over age processing, suggesting that it may “dominate” over age processing.

Given this, we might argue that sex processing is the dominant task. Following the logic of Allport et al., (1994, 2000) and Yeung and Monsell (2003), we might therefore expect that there will be higher task-switching costs when going from the age (nondominant) task to the sex (dominant) task. That is, sex processing is harder to inhibit and thus will suffer greater *task-set inertia*, while age will require more top-down control and will therefore be less susceptible to such inertia (see figure 7.1). However, the difference in dominance might not cross the necessary threshold which necessitates active inhibition of sex processing in order to engage in age processing (Monsell et al., 2000).

7.2.3 Sex and age in task-switching

There have been several investigations of task-switching between sex and age in faces. However, these have not typically been concerned with perceptual integrality, instead typically using the task as a means to investigate other ends. Accordingly, these studies may collapse details across the tasks and simply report switching costs or a lack thereof without considering the direction of the switch. For example, Marzecová et al. (2013) were the first to report such a task, carrying out a task-switching experiment which involved switching between a sex and an age discrimination tasks on faces which they named the Social Category Switching Task (SCST). This was developed to investigate differences in task switching capabilities between monolingual and bilingual populations.

The task has since been used to examine the effects of control deprivation (Bukowski et al., 2019), and changes in goal (speed or accuracy) on task switching efficiency (Castro, Bukowski, Lupiáñez & Wodniecka, 2021). To date, the original study by Marzecová et al. (2013) is the only one to attempt an explanation as to *why* there may be asymmetries in task-switching abilities between sex and age. Subsequent studies which used the SCST did not report their analyses in such a way as to investigate differences between the two tasks or in switching between them.

In the original experiment using SCST (Marzecová et al., 2013), only one stimulus from each of the four possible combinations of sex and age (young female, old female, young male, old male) was used. Stimuli were taken from the Center for Vital Longevity Face Database (Minear & Park, 2004), the stimulus set dubbed the "Lifespan" *set* in the current work. The categorisation task (sex or age) was indicated to participants by presenting a coloured frame (green or purple) around the stimulus on each trial.

The researchers (Marzecová et al., 2013) found no effect of task itself – participants were no faster or more accurate in one categorisation task compared to the other. They found a marginal effect showing faster reactions when an identical stimulus was repeated in the sex compared to the age task. They also found significant and asymmetrical switching costs for both tasks, such that there were higher switching costs when switching from the sex to the age task than vice versa. If I were to follow the logic of my own argument so far, and accept the *task-set inertia* account, this is the opposite of what I should expect.

Marzecová et al. (2013) found that the biggest switch costs were associated with trials in which the task changed but the stimulus remained the same. This was interpreted in terms of the need to update the stimulus-task set binding. The ability to do this was modulated by task, such that switching to the age task was slower than switching to the sex task when the stimulus was held constant. Because this effect was in relation to the same stimulus, it suggests that the effect is less related to the extraction of features from the stimulus, and more attributable to shifting the intentional set. It may be more difficult to adopt an age processing strategy than a sex processing strategy, or it may be more difficult to switch from the former to the latter than vice versa. This speaks to the requirement for more top-down control in the nondominant (age) task as described by Yeung and Monsell (2003).

In terms of an interpretation for the differences between tasks, Marzecová et al. (2013) appeal to the findings of Mouchetant-Rostaing and Giard (2003) – who showed using ERP and explicit sex and age categorisation tasks that greater amplitudes and longer lasting activities in occipito-parietal regions were associated with age compared to sex judgements. Mouchetant-Rostaing and Giard (2003) argued that the larger amplitudes for age than for sex judgements may reflect the "higher difficulty" of age compared to sex judgments, although the behavioural evidence tends to find that sex processing is actually more difficult than age processing (Fitousi, 2020; Quinn & Macrae, 2005; Current volume, chapters 4, 5 and 6), and Marzecová et al. (2013) found no difference in RT or accuracy between the two tasks.

Despite this, Marzecová et al. (2013) accept the logic of Mouchetant-Rostaing and Giard (2003), arguing that the reduced switch cost from age to sex may reflect greater efficiency in disengaging from the more complex task and reorientating to a simpler task – logic which defies the theory propounded by Allport et al (1994, 2000) and Monsell et al., (2000, 2003). Marzecová et al. (2013) refer to the ability to ignore something that used to be relevant as "unbinding". According to this logic, it is easier to unbind the task-set associated with Sex than Age.

Based on the consistent behavioural finding that sex processing is actually more difficult than age processing, I would argue that the depth-of-processing account (DoPA) that I described in chapter 4 can be used to explain the findings of Marzecová et al. (2013). According to this account, the processing of sex information relies upon more computationally demanding configural processing than age information – as age information can be extracted from relatively superficial and easier to process featural cues, such as skin texture and face colour (e.g. Burt & Perrett, 1995) which may change throughout the lifespan as a function of age, but may also vary as a function of health and lifestyle, for example.

By the depth of processing account, then, one can "unbind" the intentional set from the attentional set associated with the sex task without needing to rely on any more, or any different attentional information in order to conduct the age task. When unbinding from the age task, one must "go deeper" into the attentional set in order to extract the more complex information necessary to make a sex judgement. That is, going from sex to age processing requires only the uptake of shallower, less complex and less resource-intensive processing than going from age to sex processing – which requires the uptake of deeper, more complex and more resource-intensive processing.

The asymmetrical integrality noted between sex and age processing may also suggest that more age processing is carried out "en route" to sex processing than vice versa. That is, some age processing may be carried out incidentally in the carrying out of sex processing – since the sex processing occurs at a deeper level. This explains the finding that age processing appears to be contingent on sex processing (see chapter 6). This reasoning helps to explain why the sex task is typically more difficult but more “dominant” – it requires the processing of more complex information, some of which underlies the information upon which age judgements are based. That is, sex processing relies more upon the configurational properties within which the features (which age processing is more reliant on) are embedded.

7.2.4 Stimuli and types of switch cost

In the SCST as utilised in previous experiments, very few stimuli are presented. The original experiment by Marzecová et al. (2013) used only one stimulus for each of the four social categories; Bukowski et al., (2018); Kossowska, Bukowski and Czarnek (2014) and Castro et al. (2021) used two photographs per social category. In each case, the “Lifespan” face set was used. When using so few stimuli, it may not be possible to fully disentangle the effects of the social dimensions of sex and age from the effects of “identity” - that being the identity of the individuals represented in the photographs. Using one stimulus *set* may also be problematic; as it has been shown throughout the experiments reported in the current volume (chapters 4, 5 and 6) that the direction and extent of experimental effects can vary as a function of stimulus properties.

In the current experiment, participants viewed one of three sets of stimuli, and – from those – eight stimuli per social category. I tested participants on *pure* and *mixed blocks.* In pure blocks, only one task is undertaken, whereas in mixed blocks participants alternated between tasks at predictable intervals (Wasylyshyn, Verhaeghen & Sliwinski, 2011). This design permitted us to investigate two types of *switch cost,* each of which can provide an index into the processes involved in switching between reporting the age and sex of a face. *General switch costs* are the difference between performance on a given task in *mixed blocks* and performance on that task in *pure blocks. Specific switching costs* reflect the difference in performance on trials in which one has just switched to a new task (*switch trials)* compared to when they are repeating the same task (*nonswitch trials*) and are measured within *mixed blocks* (Kray & Lindenberger, 2000, 2002).

General costs are thought to be an index of the cost of maintaining and scheduling multiple task-sets in working memory (Kray & Lindenberger, 2002). It is assumed that performance in the

Nonswitch trials would be worse than performance in the Pure blocks due to the increased load on working memory associated with knowledge about the upcoming switch to another *task*. Of experimental interest is whether this effect is modulated by *set* and *task*, since this would show an asymmetry between *tasks* with relation to the performance cost associated with keeping the irrelevant *task* in working memory. If it is easier to maintain the irrelevant task in working memory, we should expect less interference from it and thus lower *General costs* compared to performance in the *pure blocks.*

Specific costs are thought to index the executive process required to deactivate the task-set relevant to the previous *task* and to activate the next one. Interference between task-sets is proposed to occur at the level of response selection, since it is a measure of the difference between two trials on the same task presented one after the other. If it is easier to inhibit one task-set and set the other up, we should see lower *specific costs* for that task compared to the other.

Investigating cognitive ageing, Kray and Lindenberger (2000) found that *general* and *specific costs* did not diminish with practice and generalised across verbal, figural and numeric stimuli – indicating that the measures reflect basic and domain-general aspects of cognitive control. They also found behavioural evidence for a dissociation between *general* and *specific costs*.

7.2.5 Predictions

I expect to find asymmetries between these two types of switching costs when switching between the two different *tasks*. More specifically, I expect that switching from sex to age will be associated with smaller switching costs than switching from age to sex because of the structural-reference (Ganel et al. 2004) and parallel-contingent (Johnstone & Downing, 2017) hypotheses, and the depth-of-processing account (DoPA) described above (chapters 4 and 6). Taking these together, sex processing is a deeper (more complex) process within which age perception processes are embedded.

It is possible that no asymmetrical *specific* switching costs occur, since my findings in the flanker task and word priming task (chapters 5 and 6) suggest that the interference between sex and age does not occur at the response selection level – which is indexed by *specific* costs in particular (Kray & Lindenberger, 2002).

Better understanding *switching costs* between these types of judgement will allow us to better understand the nature and dynamics of the interference between these dimensions when they are in direct competition for perceptual representation and in working memory, as well as allow us to determine whether the difference in dominance between the two tasks is sufficient to engender active task-set inhibition.

7.3 Methods

7.3.1 Participants

A total of 199 participants took part. 15 (7.54%) participants were excluded as outliers in terms of mean accuracy or RT (7.53%), leaving 184 participants. Following these exclusions, 0.91% of individual trials were removed from the analysis (<200/>2500ms). Of the remaining participants, 183 (99.46%) reported demographic information. 91 (49.45%) reported their sex as male, 92 (50%) as female. The mean age of participants who provided age information was 35.37 (min = 19, max = 71). Participants were paid £1.88 (Median duration = 11:42). All procedures were granted ethical approval by the Research Ethics Committee of Bangor University's School of Psychology.

SET	**N**
FACEGEN	62
MPI	62
LIFESPAN	60
TOTAL	184

Table 7.1: Number of participants per *set*

7.3.2 Stimuli

Please see the Chapter 3 for detailed information about the face stimuli. In addition to the face stimuli there were instruction screens presented to participants in the centre of the screen. These highlighted the social dimension that the participant should pay attention to (Sex or Age) and showed the possible responses and keyboard keys associated with them (see table 7.3).

7.3.3 Design

In a mixed design, *set* (Facegen, MPI or Lifespan) was a between-subjects variable and there were three within-subjects variables, *task* (Sex and Age), *trial type* (Pure, Stick and Switch) and *block type* (Pure and Mixed). The interaction of these variables varied as a function of the type of switching cost under analysis – more specific details about these analyses are provided in the results section. Each participant took part in two Pure blocks followed by two Mixed blocks: one Pure block for each of the two *tasks* and one Mixed block beginning with one or the other *task.*

The experiment was counterbalanced in terms of *block order* (see table 7.2) and *key compatibility* (see table 7.3). In terms of *block order,* Pure blocks were always presented *before* the Mixed blocks because it was assumed that performance would be better in the Pure compared to the Mixed blocks – and there was therefore no need to counterbalance them. Only the order in which each of the Pure blocks and each of the Mixed blocks were presented was counterbalanced. This yielded four possible *block orders* that the blocks could appear in (see table 7.2).

Table 7.2: The four possible *block orders* presented between participants

Order	Pure Blocks		Mixed Blocks	
	1	2	1	2
1	Age	Sex	Age	Sex
2	Age	Sex	Sex	Age
3	Sex	Age	Sex	Age
4	Sex	Age	Age	Sex

Key Compatibility refers to the compatibility of dimensional values across response patterns for each *task*. It has been noted that Young and Female on the one hand, and Old and Male on the other are compatible values (Fitousi, 2020; Kloth et al., 2015; Current volume, chapters 4, 5 and 6). For each *task*, the participant responds with two fingers on one or the other side of the keyboard, using the “A”, “D”, “J” and “L” keys to respond. Participants can therefore respond to Compatible values on either the same Side of the keyboard (Left or Right), or with the same Finger (Index or Middle). If a participant is in the Side condition of *key compatibility*, they respond on the same side of the *task*-relevant dyad for compatible values; for example, they might respond to “Young” using the “A” key (on the left of “A” and “D”) and “Female” using the “J” key (on the left of “J” and “L”) with “Young” and “Female” being compatible values. In the Finger condition of *key compatibility,* they might respond to “Young” using the middle finger of one hand and “Female” using the middle finger of the other (see figure 7.2 and table 7.3).

Counterbalancing across *tasks* for each possible combination of both Side and Finger Compatibility response patterns yields eight possible response mappings (table 3). Controlling these variables offers another opportunity to test the Stroop-like phenomena observed in previous chapters, whilst also controlling for the impact of these mappings on the task-switching effects of main interest.

Table 7.3: The four response *mappings* for each of the two levels of *key compatibility*.

Map	Compatible levels on same side of keyboard			
	Female	Male	Young	Old
1	A	D	J	L
2	D	A	L	J
3	J	L	A	D
4	L	J	D	A

	Compatible levels responded to with same finger on each hand			
	Female	Male	Young	Old
1	A	D	L	J
2	D	A	J	L
3	J	L	D	A
4	L	J	A	D

Participants were randomly assigned to one of three stimulus s*ets*, four *block orders* (see table 2) and eight *mappings* – which yielded the variable of *key congruency* (see table 3). This yielded a total of 96 experimental conditions (*set* = 3, *order* = 4, *mapping* = 8; 3x4x8 = 96).

7.3.4 Procedure

Please see the *General Methods* section of chapter 3 for the experiment-general procedure. Participants were instructed that they would be presented with faces about which they would be required to make social categorisation judgements. They were asked to their middle and index fingers of the left and right hand on the “A” and “D”, and “J” and “L” keys respectively. They were first presented with two Pure blocks, on after the other. Each block contained 32 trials and there were 32 face stimuli in each *set,* so each of those faces was shown once per block. In the Pure blocks, participants were shown the instructions for one task (2000ms) followed by the experimental trials. Each trial consists of a fixation cross (1000ms) followed by a target face (3000ms or until response). Following these blocks, participants proceeded onto the Mixed blocks, in which the task instructions changed after every two trials.

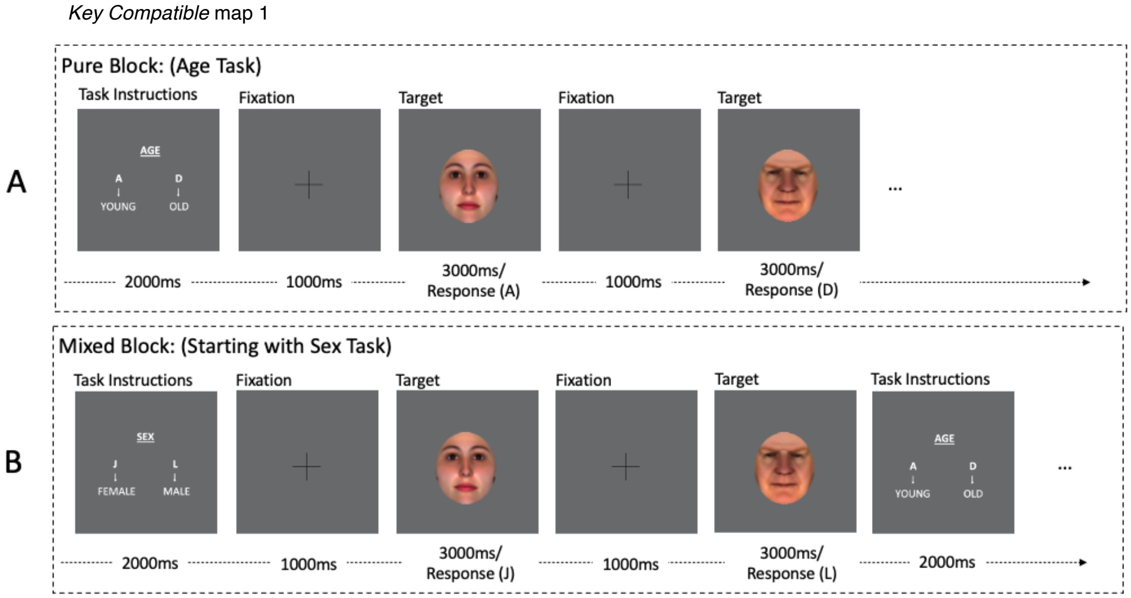


Figure 7.2: Examples of each *block type* in a *key compatibility* version of the task (where the congruent levels of Sex and Age were responded to using the same side of each of the pairs of keyboard keys). A) an example of trials in the Pure block in which the *task* is Age. B) Examples of trials in which the block is Mixed, starting with the Sex *task*.

7.4 Results

7.4.1 Control Analyses

In order to ensure that no significant degree of difference in performance could be attributed to response biases or order effects that were not of experimental interest, two separate 4x4 two-way between-subjects ANOVAs were carried out comparing performance (IES) with the control factors *order* (1, 2, 3 or 4) and *mapping* (1, 2, 3 or 4) as factors, with one ANOVA for each type of *key compatibility* (Side or Finger). There were no significant effects of *order, mapping* or *compatibility* and no interactions (all $F < 1.1$ and $p > 0.32$). This meant that there were no significant order effects or confounding response biases.

To check that there was no difference between the types of *key compatibility*, which would be indicative of an experimentally interesting response bias, I carried out an independent t-test comparing performance between cases where participants responded on the same side of the keyboard (Side compatible) or with the same finger on each hand (Finger compatible). This test was insignificant (t(365) = 1.14, p = 0.26; Side M = 976.59 (2.46), Finger M = 1004.79 (2.49)). These results increases confidence that *key compatibility* is not a major confound in the experimental analyses.

Next, I carried out an analysis to establish baseline performance for each *task* and *set* so as to inform the interpretation of the main experimental analyses. I collapsed IES scores on these factors across *block type* and *block order* and used a two-way mixed subjects ANOVA with *task* as a within-subjects factor and *set* as a between-subjects factor. This analysis revealed a significant main effect of *set* (F(2,181) = 5.14, p = 0.007, $h^2 = 0.05$). Fisher's LSD (87.25) showed that performance was not significantly different between the Facegen (M = 993.93 (34.1)) and MPI (M = 974.06 (28.1)) *sets*, but that performance was significantly better for both of these *sets* than for the Lifespan *set* (M = 1106.19 (31.32)).

There was also a significant interaction between *task* and *set* (F(2,181) = 8.62, $p < 0.001$, $h^2 = 0.006$). Follow-up t-tests showed a significant difference for the Facegen *set* such that performance was significantly better in the Sex *task* (t(185) = 3.16, $p < 0.001$, d = 0.16; Sex M = 968.11 (36.09), Age M = 1019.74 (33.62)), and was significantly better for the Age *task* for the MPI *set* (t(185) = 2.88, $p < 0.004$, d = 0.16, Sex M = 997.16 (29.72), Age M = 950.95 (28.93)). There was no significant difference between *tasks* for the Lifespan *set* (p = 0.29) (see figure 7.3)

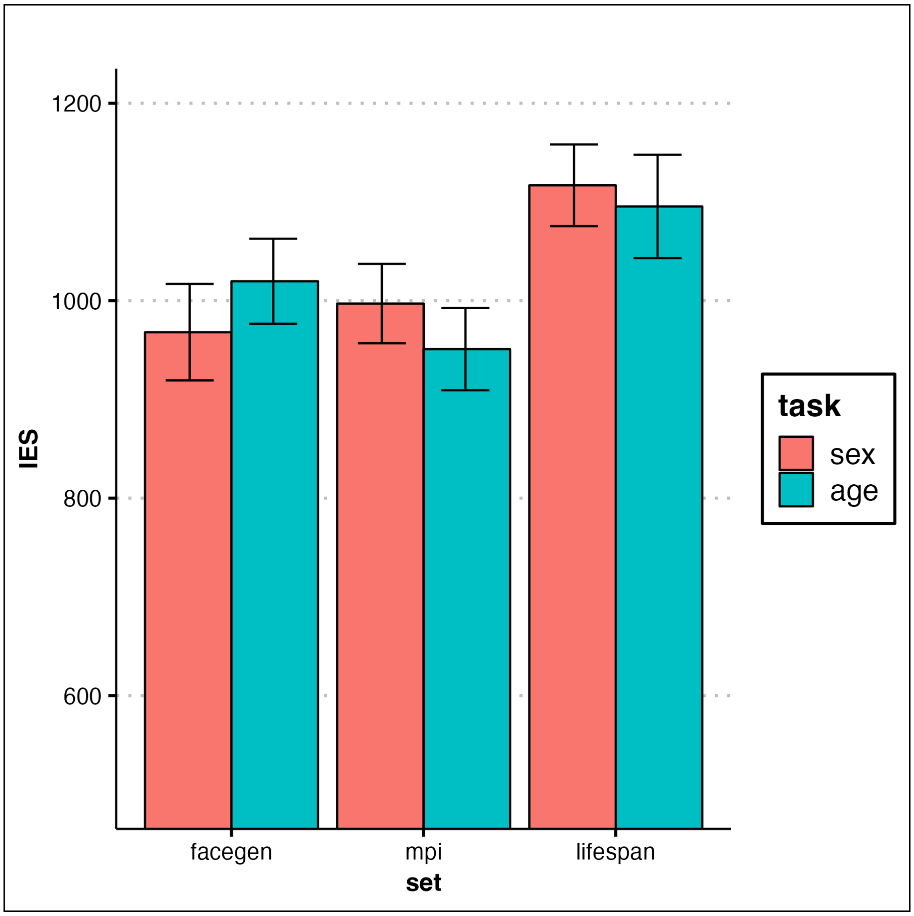


Figure 7.3: Mean IES for each *set* split by *task.*

7.4.2 General switch cost

For this analysis, a difference score was calculated by subtracting performance on the Pure blocks from performance on the non-switch trials in the Mixed blocks – that is, on the trials in which there was no switch but the continuation of the same *task*. A positive difference score would indicate that performance was better in the Pure blocks compared to the non-switch trials of the Mixed blocks. To investigate the general switching cost and its possible modulation by *set* and *task,* difference scores were submitted to a 3x2 two-way mixed-subjects ANOVA with *set* and task as factors*.* Participants whose absolute mean difference score – averaged over both *tasks* - fell outwith 2.5 SDs of the absolute mean difference score - averaged across *task* and *set* - were excluded from the general switch costs analysis (n = 6).

This analysis showed an overall general switching cost, interpreted as a y-intercept significantly different from zero (F(1,175) = 26.28, p < 0.001, $h^2 =$ 0.075; M = 58.14 (17.29)). Performance was better in the Pure blocks (M = 920.97 (15.72)) than in the Nonswitch trials of the Mixed blocks (M = 1051.66 (20.36)). There was no significant effect of *set* (F(2,169) = 0.52, p = 0.59) but there was a significant effect of *task* such that the general switching cost was higher for the Age *task* than the Sex *task* (F(1,175) = 14.01, p < 0.001, $h^2 =$ 0.036; Sex = 20.57 (15.52), Age M = 95.71 (18.24)). There was a significant interaction between *set* and *task* (F(2,175) = 35.03, p < 0.001, $h^2 =$ 0.155).

Follow-up analyses revealed that the general switching cost for the Facegen *set* was significantly larger for the Sex *task* than the *Age* task (t(114) = 5.04, p < 0.001, d = 0.91; Sex M = 118.11 (20.93), Age M = -50.73 (26.18)). For the MPI *set* the cost was bigger for the Age *task* (t(110) = 5.21, p < 0.001, d = 0.94; Sex M = -30.34 (20.83), Age M = 151.49 (27.99)). Similarly, for the Lifespan *set* the cost was bigger for the Age *task* (t(109) = 4.86, p < 0.001, d = 0.92; Sex M -30.24 (33.78), Age = 194.49 (31.55)) (see figure 7.4).

Interestingly, it appears that – for each *set* – there is a general switch cost for one *task* and a small switching *gain* for the other *task* – suggesting that keeping a certain irrelevant task-set in working memory actually slightly facilitates performance on the relevant task. In a post-hoc analysis, the *tasks* of each *set* (Age for Facegen and Sex for MPI and Lifespan) which displayed a negative difference score (switching gain) were collapsed across *task* and *set* and subjected to a one-sample t-test to determine if – as a whole – these switching gains were significantly different from zero. This test was significant (t(171) = 2.64, p = 0.008, d = 0.2).

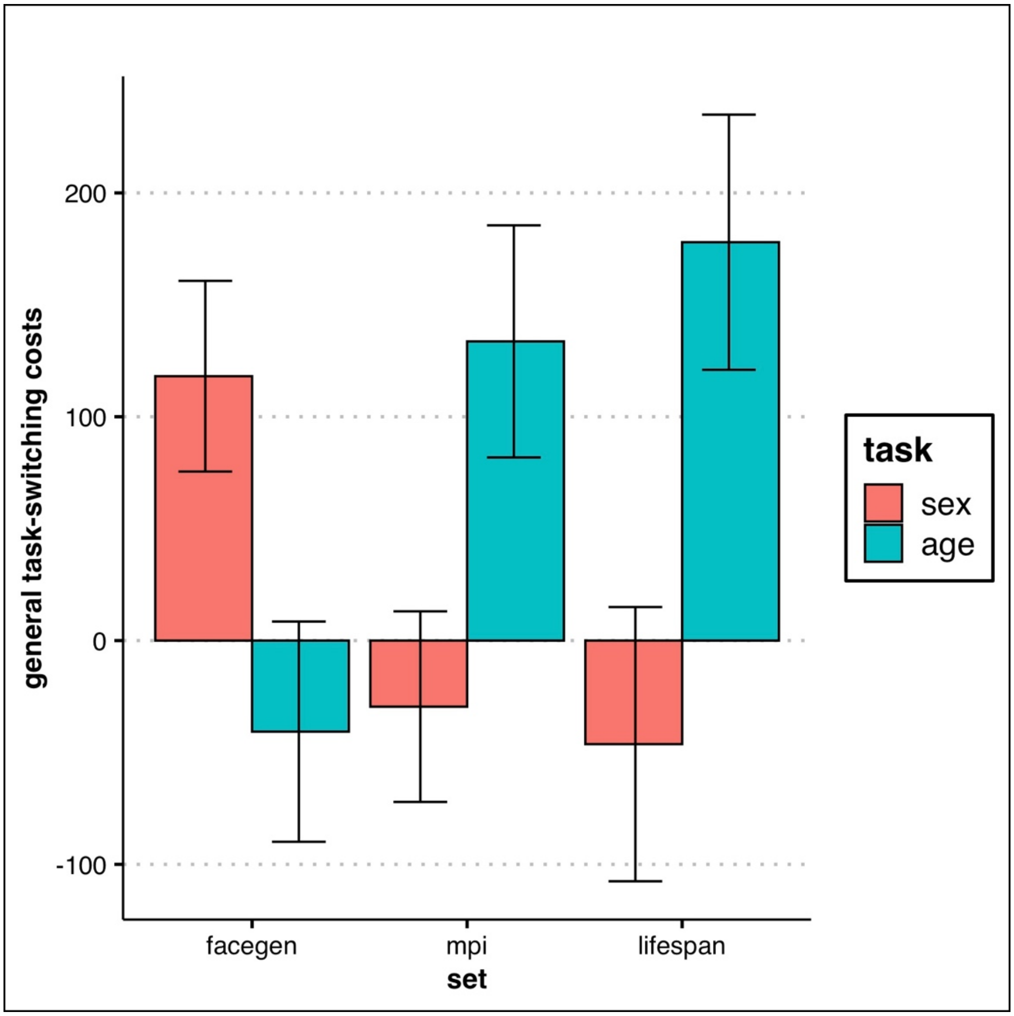


Figure 7.4: General task-switching costs for each *set*, split by *task*. Positive values indicate better performance in pure compared to mixed blocks, indicating a relative cost of performing a given task in the context of switching blocks vs pure blocks.

7.4.3 Specific Switch Costs

Specific switch costs relate to the difference between Switch trials and Nonswitch trials in the Mixed blocks, and are a measure of the cost of having just switched to a new task compared to performance on a trial which is a continuation of the same task. For this analysis, the difference score was generated by subtracting performance on the Stick trials from performance on the Switch trials – that is, for each pair of trials, subtracting the second trial from the first trial. It was assumed that performance would be worse on the Switch trials compared to the Stick trials on account of the difficulty associated with switching response patterns. The resultant difference score should therefore be positive. Of interest was whether this switching cost was modulated by *task* and *set,* indicating an asymmetry between switching response patterns between the easier and more difficult *task*.

To investigate the specific switching cost and its possible modulation by *set* and *task,* difference scores were submitted to a 3x2 two-way mixed-subjects ANOVA with *set* as a between-subjects factor and *task* as a within-subjects factor. Participants whose absolute mean difference score fell outwith 2.5 SDs of the grand absolute mean difference score were excluded from the analysis ($n = 2$).

This analysis showed an overall specific switching cost, interpreted as a y-intercept significantly different from zero ($F(1,179) = 155.27$, $p < 0.001$, $h^2 = 0.364$; $M = 148.75$ (14.75)), indicating that performance was generally better in the Nonswitch trials compared to the Switch trials. This effect was not modulated by *set* ($p = 0.36$), *task* ($p = 0.16$) or an interaction between *task* and *set* ($p = 0.18$), meaning that there were no significant differences in terms of the way that the specific switching cost manifested across *sets* and *tasks*.

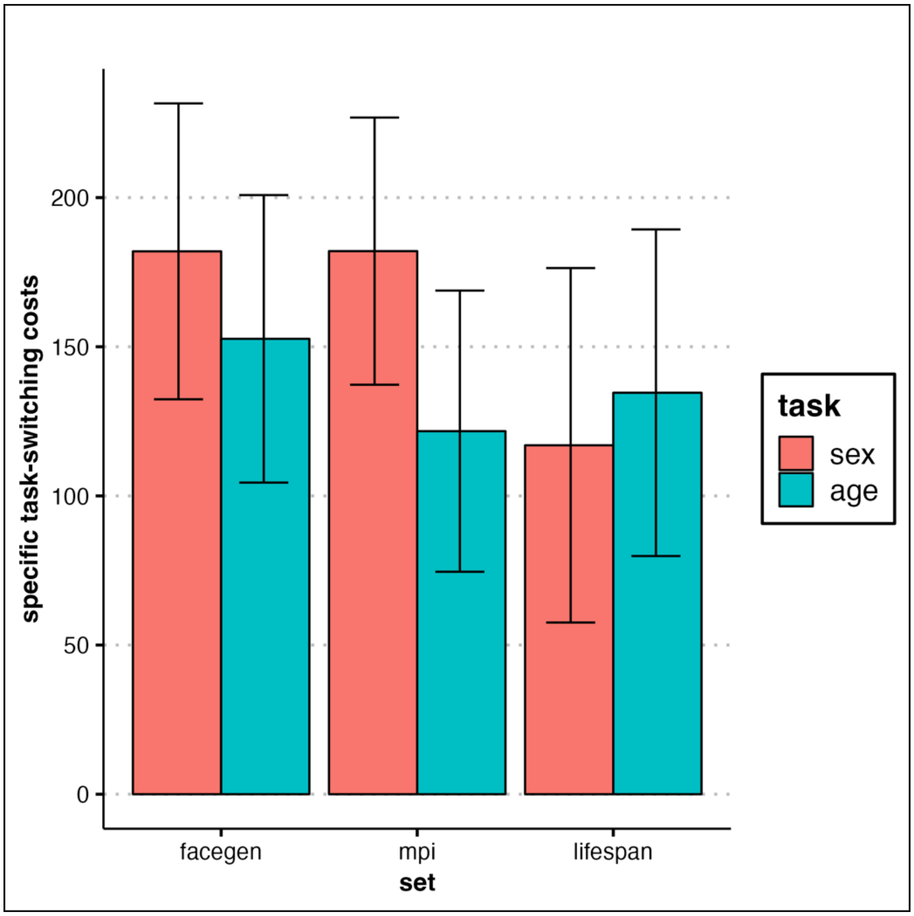


Figure 7.5: Specific switch costs for each *set*, split by *task*.

7.4.4 Post-hoc analysis

The results are complicated by the difference in task-difficulty – and also because the difference in task-difficulty changes between *sets*. In order to explore this difference in task-difficulty more deeply and to facilitate the interpretation of the results, I looked at performance in the *pure blocks* more closely. Specifically, I wanted to know if there was a difference in terms of how performance on each task improved as the block progressed. This information might tell us if one *task* is easier to "automate" than the other. Increased speed and accuracy of processing (as captured by IES) can be thought of as a marker of automation with practice (e.g. Jansma, Ramsey, Slagter & Kahn, 2001). Determining whether one or the other *task* is more easily automated could help to constrain the interpretation of the other results.

To investigate this, I took the mean RT on accurate trials (IES) for the first and second half of each *pure block* for each *task*. This gave me a factor called *half* with a value for the first half and the second half of the *pure blocks* – or trials 1-16 and 17-32 respectively. I collapsed this measure across both *pure blocks* (one for each *task*) regardless of which order the participant completed them, since the order of Pure *blocks* was counter-balanced across participants. I conducted a paired samples t-test comparing performance on the first *half* of each *task*. This analysis was not significant ($t(183) = 0.91$, $p = 0.36$) indicating that performance in Sex and Age discrimination was not significantly different in the first half of Pure blocks for those *tasks*.

I then created a difference score by subtracting performance in the second half of the blocks for each *task* from the first, leaving a score called the “automation effect”. A positive difference score indicates better performance in the second half of the block compared to the first half. Participants whose absolute mean difference score fell outwith 2.5 SDs of the grand absolute mean difference score were excluded from the analysis ($n = 2$). I submitted the difference scores to a two-way mixed effects ANOVA, with *set* as a between-subjects factor *task* as a within-subjects factor.

This analysis revealed a significant automation effect, interpreted as a y-intercept significantly different from zero ($F(1,177) = 118.25$, $p < 0.001$, $h^2 = 0.26$, $M = 84.42$ (10.74)), indicating that performance was significantly better in the second than the first *half* of Pure blocks. There was a main effect of *task* which indicated that the automation effect was larger in the Age compared to the Sex *task* ($F(1,177) = 6.43$, $p = 0.01$, $h^2 = 0.017$, Sex $M = 65.73$ (10.65), Age $M = 103.1$ (10.68)) (see figure 7.6). One-sample t-tests for each *task* indicated that there was a significant automation score for both the Sex ($t(179) = 6.17$, $p < 0.001$, $d = 0.46$) and the Age *tasks* ($t(179) = 9.65$, $p < 0.001$, $d = 0.72$).

The lack of an effect of *set* ($p = 0.54$) and of an interaction between *task* and *set* ($p = 0.92$) suggests that automation did not significantly vary as a function of the stimulus properties, and that the pattern of increased automation in the Age relative to the Sex *task* was broadly consistent between *sets*.

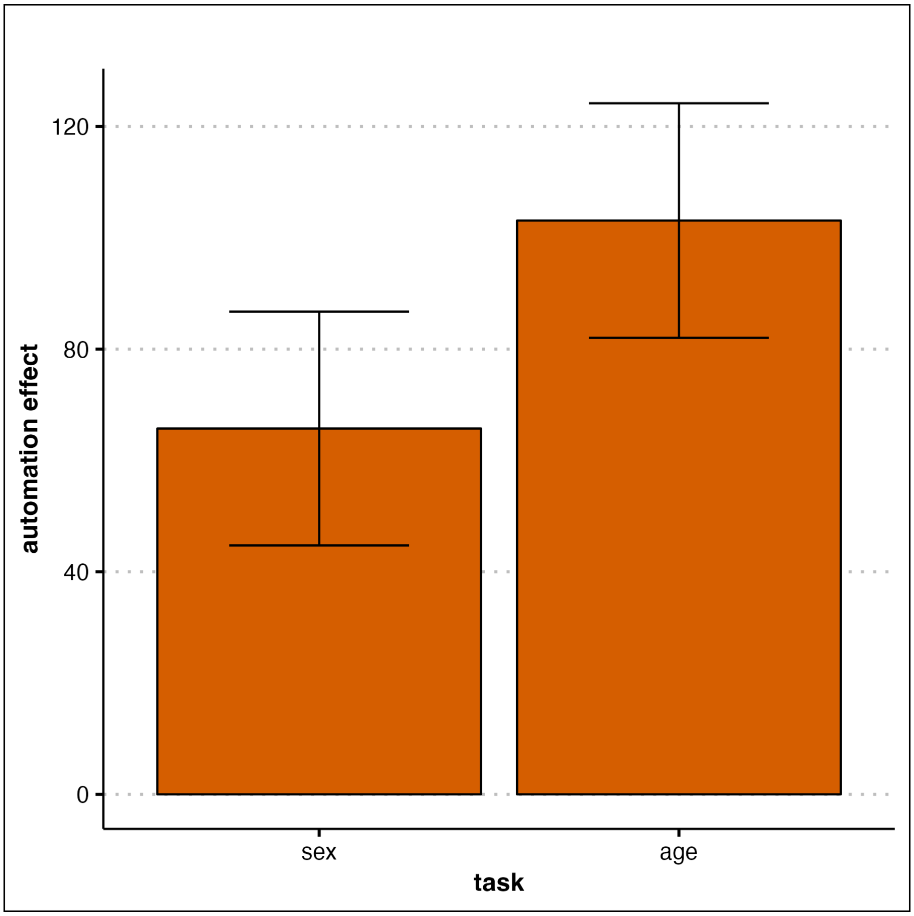


Figure 7.6: Automation effect for the Pure *blocks* split by *task* reveals that judgements become more efficient for both *tasks* over the course of a block, but that performance on the Age *task* improves more than it does for the Sex *task* – regardless of task-difficulty relative to the *set*.

7.4.5 Stroop effects

Difference scores were calculated by subtracting the Compatible from the Incompatible configurations of dimensional values. Stroop effects and their possible modulation by *task* and by stimulus *set* were assessed by submitting difference scores to a 2x3 two-way between-subjects ANOVA with *task* and *set* as factors. Prior to analysis, participants whose difference score (averaged across *tasks*) fell outwith 2.5 SDs of the mean difference score were excluded from the Stroop analysis (n = 3).

This analysis showed a significant, positive Stroop effect, interpreted as a y-intercept significantly different from zero (F(1,178) = 57.73, p < 0.001, h^2 = 0.165; M = 60.37 (10.6), This effect was modulated by *task* such that the Stroop effect was larger for the Age *task* than for the Sex *task* (F(1,178) = 7.52, p = 0.01, h² = 0.016; Sex M = 43.32 (10.38), Age M = 77.41 (10.45)), and by *set* (F(2,178) = 11, p < 0.001, h² = 0.07; Facegen M = 111.13 (12.23); MPI = 37.14 (13.55); Lifespan M = 30.42 (14.98)). FLSD (37.86) showed that the significant effect of *set* was characterised by a significantly higher Stroop effect for the Facegen *set* compared to both the MPI and the Lifespan *set* – which, in turn, were not significantly different from one another.

These main effects were qualified by a significant interaction between *task* and *set* (F(2,178) = 3.78, p = 0.02, h² = 0.016). Follow-up analyses revealed a significant difference for the Facegen *set* such that the Stroop effect was larger in the Age compared to the Sex *task* (t(103) = 3.19, p = 0.002, d = 0.57, Sex M = 75.67 (11.96), Age M = 146.58 (18.73)). The difference between *tasks* was not significant for the MPI (p = 0.63) or the Lifespan *sets* (0.13)

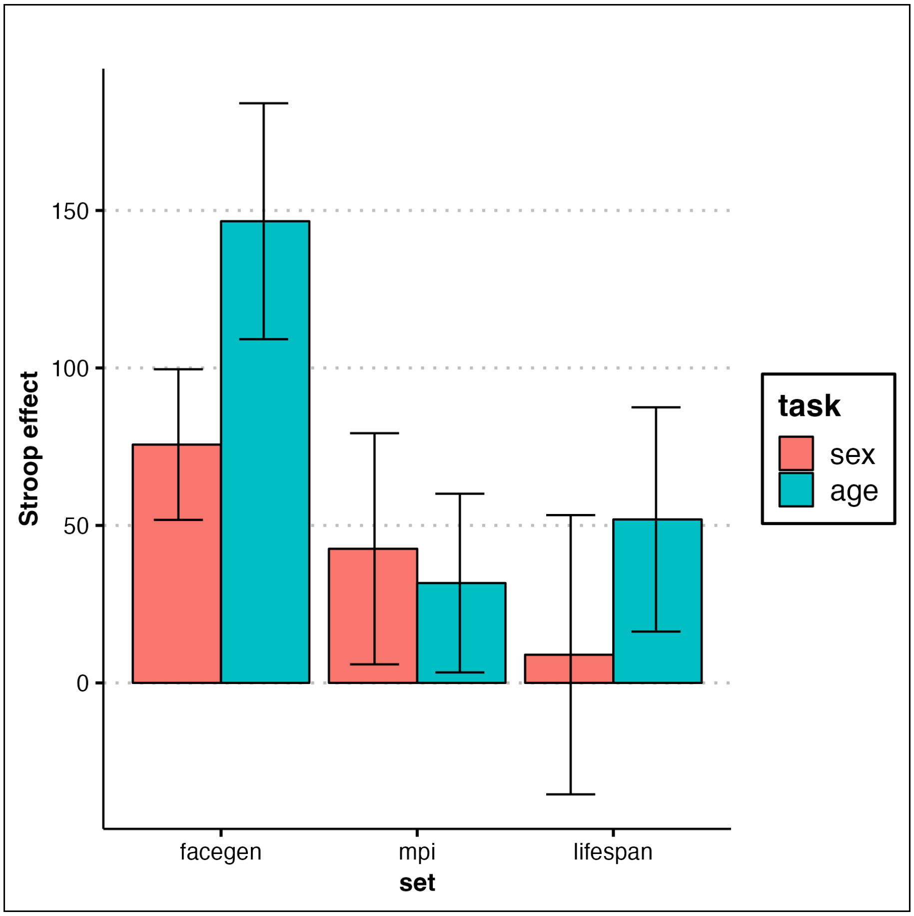


Figure 7.7: Stroop Effect for each *set* split by *task*. Positive values indicate better performance on Compatible compared to the Incompatible stimuli.

7.5 Discussion

General and specific *switching costs* were observed when switching between Sex and Age discrimination *tasks*. The general *cost* suggests that the *task-sets* associated with sex and age processing interfere with one another in working memory – that the maintenance and scheduling of these two *task-sets* negatively affects performance relative to when only having to maintain and schedule a single *task-set.* The specific *cost* suggests that switching between *task-sets* incurs a cost compared to sticking with the same *task-set*. An asymmetry was found for the general, but not the specific switching costs. This asymmetry was such the task-set for the Age *task* interfered with Sex processing more than vice versa for the Facegen *set,* and the task-set for the Sex *task* interfered with processing for the Age *task* more than vice versa for the MPI and Lifespan *sets.*

I also found more evidence for Stroop compatibility between age and sex, and the post-hoc "automation effect" analysis suggests that the processing of age can be more easily automated than the processing of sex (with automation being described in terms of increased speed and accuracy of processing with practice (e.g. Jansma et al. 2001)). The finding that age is more easily automated than sex is in line with previous findings (Wiese et al., 2008, 2012), and supports the idea that age perception is relatively more reliant upon bottom-up processing strategies. These two findings, as well as the differences in baseline task-difficulty across *sets*, complicate but inform the interpretation of the switching cost analyses. I will begin by exploring the baseline, Stroop and automation analyses and then the switching cost analyses in relation to them. I will then offer an interpretation before concluding.

7.5.1 Baseline, Stroop and automation analyses

Task "dominance" is regarded as a central factor in determining asymmetric switching costs (Allport et al., 1994; Monsell & Yeung, 2003). Specifically, that there is a higher cost when going from the "nondominant" to the "dominant" (i.e. pre-potent) task than vice versa. In the current results, the pattern of differences in task-difficulty and the direction of the asymmetry of the general switching costs is generally consistent between *sets* – with greater switching costs for the "dominant" *task.* This suggests a prominent role of task-difficulty in the emergence of general switching costs, such that keeping the harder *task* in mind leads to decrements in performance when conducting that task, in line with the account of Allport et al. (1994) and Monsell and Yeung (2003). This account, however, does not seem to tell the full story.

Looking at the Facegen *set*, sex is significantly easier than age (figure 7.3), but age remains – like with the other two *sets* – a more easily automatable *task* (figure 7.6), casting doubt as to which *task* is “dominant”. The Facegen *set* also revealed the greatest Stroop effect (figure 7.7), and was the only one for which the Stroop effect was asymmetrical between *tasks* – with a larger Stroop effect in the age *task* – with the age task being the more difficult but more easily automatable *task*.

For the Lifespan *set*, there was no significant difference between *tasks* in terms of difficulty (figure 7.3), a difference in terms of which *task* was more easily automated (figure 7.6), and a small Stroop effect (figure 7.7). The Lifespan *set* was also the *set* for which there was the greatest general switching cost and the largest asymmetry in that cost between *tasks*. If asymmetric switching costs were characterised by differences in task-dominance (defined in terms of ease, automatability and being “better learned” (Gilbert & Shallice, 2002)), we would expect to see smaller switching costs, not larger.

It is also conspicuous that the pattern of automaticity of each respective *task* does not map neatly onto the pattern of task-difficulty for each *set*: more specifically, age processing is more automatic but also more difficult for the Facegen *set*. This may indicate that some of the observed interference between these two dimensions is occurring at higher, more semantic and less visual levels. Taken together, such discrepancies seem to indicate that stimulus properties tell some, but not all of the story.

There was a significant Stroop effect for each *set,* with a difference between *tasks* for the Facegen *set*, such that there was a larger effect for the more difficult Age *task*. As to why there was a difference for this *set* but not the others I can only speculate. It may have to do with the recruitment of task-irrelevant stimulus properties in difficult judgements with unfamiliar stimuli (Facegen faces). This would explain why the Stroop effect was larger for the Age than the Sex *task* with the Facegen *set* as well as why there was a comparatively larger Stroop *effect* for the

Facegen *set* compared to the others, but would not explain the lack of difference between *tasks* for the photographic *sets*.

7.5.2 General switching costs

There were asymmetrical general switch costs (that is, the cost of nonswitch trials in the Mixed blocks compared to the Pure blocks for a given task) such that, for the photographic *sets* (MPI and Lifespan), there was a significant switching cost for the Age but not the Sex *task* – which was the same finding reported by Marzecová et al. (2013). In contrast, for the artificial *set* (Facegen), the order was reversed - with switching costs for the Sex but not the Age *task*.

The greater general switching cost for each *set* occurred within the easier task. The increased working memory load associated with the upcoming switch to the harder *task* inhibits performance more than the working memory load associated with switching to the easier *task.* Keeping the harder *task* in mind whilst performing the easier *task* leads to greater decrements in performance than keeping the easier *task* in mind whilst performing the harder *task*. The modulation of this effect by *task* and *set* suggests that this effect is driven by stimulus features rather than at more semantic levels, however the fact that the largest asymmetry in general costs is for the Lifespan *set* - for which there is the smallest difference in task-difficulty - muddies this interpretation.

There are a number of potential interpretations for the general switching cost findings. The first follows the logic of Allport et al. (1994) that asymmetric task-switching costs are typically larger going from the dominant to the non-dominant task. Taken in tandem with the task-difficulty and automation results, this finding suggests that age processing may be the "dominant" task for the photographic *sets*, and that sex processing may be dominant for the Facegen *set.* However the finding that age processing was more easily automated even for the Facegen *set* does not fit with this interpretation.

The increased interference in the dominant *task* may occur because of *carry-over inhibition* (Monsell et al. 2000), which results from the interfering and distracting elements of the dominant *task-set* being inhibited when carrying out the nondominant *task* and then persisting when performing the dominant *task* again*.* The inhibition of the dominant task-set allows the observer to ignore the elements of the stimulus relevant to that *task* and focus on the relevant *task-set* when performing the nondominant *task.* This explains the facilitated nondominant processing following dominant processing noted for each *set* (see figure 7.4). Although this switching gain was insignificant for each *set*, the switching gain was significantly different from zero when collapsed across *sets*, and occurred for the more difficult *task* in each *set*.

A simpler explanation is that (at least for the photographic sets) the switching costs for the dominant (Age) *task* might be characterised by the lack of an opportunity to automate the task. In the Pure blocks, we see that age processing improves in efficiency over time more so than sex processing. Since general switching costs capture the difference between performing a *task* in Pure compared to Mixed blocks, the fact that age suffers a greater cost than sex (in the photographic *sets*) is perhaps not a surprise. The age *task* being easier than the sex *task* for these sets may be due to the capacity to automate age processing more readily.

This automation process may be disrupted by the need to switch to another *task-set* (for sex processing), and this disruption may manifest as a general switch cost. That is, in the Pure blocks, automation may proceed, but in Mixed blocks it is constantly being interrupted. The above interpretation may account for the photographic *sets* but not the Facegen *set*, where age processing was more easily automated but more difficult, and for which there were general switching costs for the sex but not the age *task*. If age were more easily automated, and the general switching cost was an index of the interruption of this automation, we would expect to find general switching costs for the Age *task* of the Facegen *set.*

Another interpretation rests on the assumption (outlined above in chapter 4) that, for photographs, sex processing is a more complex and computationally resource-intensive process than age processing. It may be for this simple reason that sex interferes with age more in the photographic *sets* – the task-set is simply more complex and resource-intensive, and so keeping it in mind impairs performance more. This effect may be reversed for the Facegen *set* because of the caricatured nature of the sex differences between stimuli – perhaps permitting the adoption of a feature-driven discrimination strategy – and the relative unfamiliarity of the stimuli, perhaps leading to difficulties in adopting such a strategy for age processing.

As discussed, the general switching costs do not align perfectly with the patterns of task-difficulty and automation, but do appear to be highly mitigated by such patterns. It is therefore reasonable to assume that the asymmetry is arising at a perceptual level, and not at the level of response preparation and selection. This is further illustrated by the specific switching cost findings.

7.5.3 Specific switching costs

It was more difficult to carry out sex or age processing when it was necessary to retain the *task-set* associated with the Irrelevant dimension in working memory than it was when it was only necessary to maintain the *task-set* associated with the relevant *task*. Specific switch costs are supposed to index how well the cognitive system establishes new stimulus-response task-settings and deactivates previous ones (Kray & Lindenberger, 2002). This suggests that sex and age interfere with one another in this regard, however there is no asymmetry noted at this level of processing.

This suggests that the asymmetric integrality between sex and age does not manifest at the response level – in keeping with the findings reported above (chapters 4, 5 & 6). Although the task-sets interfere at the response level, they do not do so asymmetrically – it is just as easy to go from responding in terms of sex to age than from age to sex. The specific switching cost did not differ between sets, further indicating that response-level interference does not appear to vary as a function of specific stimulus properties.

7.5.4 Interpretation

The use of three stimulus *sets* in the current experiment as well as the post-hoc automation analysis has revealed that task-difficulty or task-dominance plays a role in the emergence of switching-costs, but cannot account for the results completely. The depth-of-processing account (DoPA) described above (see section 4.6) offers a framework for diving more deeply into *why* certain tasks may be more or less difficult or dominant, and may provide resolutions for some of the apparent discrepancies between *sets* and *tasks*.

Age discrimination appears to be more easily automated than sex discrimination – regardless of differences in task-difficulty between *sets*. Within the DoPA, this is because age discrimination relies on more feature-based, unstable, deformable, simple templates relative to sex discrimination. This means that age processing can be more easily carried out using features in a bottom-up manner (and, therefore, a faster and more automatic manner (see Theeuwes, 2018)), and suggests that the stimulus features can be more easily mapped to internal representations – with those representations being more malleable, meeting the stimulus properties half-way, as it were.

As the block progresses, the process of matching features within the stimulus *set* to representational templates becomes more efficient. This occurs to a lesser extent in sex discrimination, which relies upon more configuration-based, stable and complex templates which are less deformable. Since the templates are more stable and less deformable, incoming information about faces must in-effect match to the templates more strictly before a sex discrimination judgement can be made. This renders the process less automatable – each face must be processed to a deeper and more complex level before an accurate judgement can be made, unlike in age-discrimination where simpler, more surface level or "shallower" features can be relied upon.

This explains why age is still more easily automated within the Facegen *set* despite age processing being more difficult. The stimuli are more unfamiliar than photographs which may make the extraction of age information more difficult, but despite this the observer will become more familiar with the stimulus *set* over time and can automate the process because the relatively unstable age-processing templates deform to match the incoming stimulus features.

It was suggested above (chapter 4) that sex processing in the Facegen *set* may be more easily carried out using a bottom-up strategy relative to the other *sets*, since the features (argued to be processed from the bottom-up) which characterise the sex of the Facegen stimuli may be more caricatured than those characterising photographs of real faces – rendering feature-based processing a more tenable strategy. Similarly, configural templates operating from the top-down may not be as useful for discriminating between the sexes of unfamiliar stimuli (such as Facegen faces), because those templates have been developed over time through exposure to real faces, and so a strategy based on bottom-up processing may be more advantageous.

It may be that – for photographs of real people – age processing is easier to automate due to a tendency to process age according to a bottom-up, feature-based processing strategy - although the features which characterise age discriminations may vary as a function of sex. This fact may make successful age processing possible with lower levels of volitional, top-down control than sex processing, with sex discrimination potentially being slower as the result of a tendency to rely on top-down processing strategies. Age processing may occur more automatically because it relies upon more easily accessible and salient visual features than does sex processing. Sex processing may require more top-down control and may take place higher, further or deeper into the ventral processing stream – relying more upon semantic and less upon sensory information compared to age processing.

7.6 Conclusions

The findings of the current experiment suggest that maintaining cognitive and perceptual *task-sets* in relation to both sex and age processing causes interference relative to when one must only maintain a *task-set* associated with one of these dimensions. This interference appears to occur at perceptual levels and at levels of response selection, however there are asymmetries at the sensory and/or semantic (perceptual) levels, but not at the response level.

In Ravizza and Carter’s (2008) terms, this means that there are asymmetrical perceptual shifting costs, but not asymmetrical rule shifting costs. Asymmetric switching costs arise when switching attentional but not intentional sets (Rushworth et al., 2005). That is, switching between age and sex processing seems to depend to a large extent on the stimulus features – it is more of a perceptual than a cognitive problem and depends upon the faces in need of processing.

The current research is evidence that the task-switching paradigm can be meaningfully brought to bear on problems of perceptual integrality in face processing. More dimensions should be tested using this paradigm. There are two outstanding research questions. The first pertains to the finding that automation may play a role in asymmetric integrality; the second relates to the locus of that asymmetric integrality in the visual processing stream. The current work suggests that asymmetric interference does not occur at response levels, but at earlier sensory and/or semantic levels. More work is needed in order to disambiguate and elucidate these levels of processing and their relation to asymmetric integrality – suggesting that some dimensions may be processed at higher or deeper levels, and others at lower or shallower levels.

Chapter 8: General Discussion

The question is not simple, neither can our answers be.

Garner - 1974

Most popular models of human face perception have a dual-route structure (e.g. Bartlett et al. 2003; Bruce & Young, 1986; Duchaine & Yovel, 2015; Haxby et al., 2000, 2002), supposing that that there is a major structural or functional distinction in the processing of different dimensions of the face, mirroring such a distinction in the neurological structures of the brain. Tenets of these models have been variously supported and challenged through research into the integrality of face processing, much of which has highlighted the complex interference that emerges in face perception when the processing of multiple dimensions is required, rather than just the processing of one dimension in isolation.

On top of this, research in the general area of social vision has revealed that specific values within social dimensions (such as the value of “male” within the dimension of “sex”) interact with the values of other dimensions (such as the value of “old” within the dimension of “age”, (see Kloth et al., 2015)). It has been suggested that such interactions may be characterised by phenotypic congruencies in the stimulus structure between values, or by a top-down, semantic component to face processing – such as expectations about differences in the emotionally expressive tendencies of the sexes (Becker, 2007). Such top-down processing components in face processing are not described by any of the dual-route models, but are captured by the Freeman and Ambady (2011) model that has emerged out of the social vision perspective.

The empirical work presented in this thesis aims to add to previous research by exploring the relatively under-investigated relationship between sex and age in face processing, analysing the relationship between these two dimensions and their values using a variety of well-used behavioural paradigms in cognitive psychology. In this, I took a converging operations approach towards better understanding the perceptual integration between these dimensions and their values, the stages in the perceptual process at which integration emerges, and which perceptual mechanisms most likely account for it.

In the current chapter, I first discuss the importance of stimulus characteristics in face perception research. I then briefly summarise the main findings from chapters 4, 5, 6 and 7. I then discuss the Stroop findings across each of the four chapters, carrying out an “omnibus Stroop analysis” and an “omnibus task-difficulty analysis” (analysing differences in the IES of each *task* for each *set*) by combining data across experiments. I relate the experimental effects and Stroop effects to baseline task-difficulty where appropriate. Following this, I outline the major overall findings of the current work, compare them against existing models of face processing, and provide a holistic interpretation – suggesting avenues of future research. I discuss methodological issues as I go along.

8.1 The use of multiple stimulus *sets*

There are numerous reasons why the use of multiple stimulus sets was beneficial in the current case. First, confidence in observed effects could be higher when observed across multiple, independently tested sets, and could be lowered when effects were not observed across multiple sets – using multiple *sets* can in this way be thought of as a type of internal replication. Second, using multiple *sets* permits inferences to be made about the rough stage of the perceptual hierarchy – ranging from the lowest sensory stage to the highest semantic stage – at which interference occurred.

My logic is that if there were significant effects which differed between *sets*, it could be assumed that the observed effects were modulated by physical stimulus properties and thus that the interference characterising the effect occurred at a lower, more sensory level. If, on the other hand, the effect was consistent across *sets*, then the effect does not seem to be modulated by the physical properties and thus could be occurring at a higher, more semantic level of perceptual processing.

A third reason why it was important to use multiple sets was that it permitted comparison between artificial and photographic images. In general, the effects in the photographic *sets* differed significantly from those in the artificial *set*. This suggests that artificial stimuli are perceived in a different way, or through the use of different perceptual strategies than photographic *sets*. That said, there were two photographic *sets* and one artificial *set*. If I were to use an additional artificial *set* and find the same results as were observed with the Facegen *set* across experiments, then the hypothesis could be tested more directly.

The fact that two of three *sets* typically displayed more similar results could easily skew the interpretation of the results –I have tried to be mindful and conservative in my interpretations. Using three *sets* instead of one has certainly made interpretation more difficult, however it has perhaps made it more reliable. If one were to take each of the three *sets* used in the current experiments and interpret the results based on those *sets* in isolation, one may well end up reaching three separate sets of conclusions.

The current results highlight that the set of stimuli used in social perception research can have a material impact on the outcome of experiments. Also, the findings suggest that it may not be reasonable to assume that findings observed using artificially rendered stimuli can be generalised to images of real human faces. The general finding of differences in perceptual effects between these *sets* points to the sensitivity of the face processing system to visual cues.

It is quite common in the literature to base social cognition experiments on faces extracted from a single database, or to mix stimuli from numerous databases. I would argue that this may not lead to as ecologically valid or generalisable results as testing hypotheses using multiple stimulus sets independently. In the current experiments, *set* was always a between-subjects factor. Future research may use it as a within-subjects variable, more directly and powerfully testing the hypothesis that results might be modulated by *set*.

8.2 Summary of main findings

In chapter 4, I investigated the integrality of sex and age using the Garner speeded classification task. Bidirectional Garner interference indicated that variations in the task-irrelevant dimension negatively affected performance compared to when that dimension was kept constant. The effect was asymmetrical such that irrelevant variation in sex affected age discrimination more so than vice versa.

This pattern of asymmetrical integrality held across all three *sets* despite differences between the *sets* in terms of task-difficulty – suggesting that the Garner effect manifests at a later, post-sensory, representational or semantic level. This was in line with the findings of Boenke et al. (2007) who showed an earlier time-course and differential ERP correlates for Stroop compared to Garner effects. In line with previous findings (e.g. Atkinson et al., 2005), the results constitute evidence against the speed-of-processing account (SoPA) (Graham & LaBar, 2007; Melara & Mounts, 1993). This is because the asymmetrical Garner effect was homogenous across *sets* despite differences in baseline task-difficulty, and that – for the photographic *sets*, the less discriminable dimension interfered more with the more discriminable dimension than vice versa. Further, the results indicated that the Garner effect must not be occurring at the level of response selection – since *task* was a between-subjects variable.

In chapter 5, I used an Eriksen flanker task to investigate whether sex and age information from irrelevant faces affected the processing of those dimensions in a target face – whether sex and age information was perceptually integrated *between* faces. I found no significant evidence for such an effect, although the insignificant effects which were present varied as a function of *set,* suggesting that the physical nature of the target and the distractors may affect interference between them. The flanker effect is typically thought to index response level interference, and so the lack of a flanker effect can be interpreted as tentative evidence for the lack of interference between sex and age at the response level. Methodological issues may have contributed to the absence of a flanker effect.

In chapter 6, I used a face priming and a word priming task to investigate the effects of task-irrelevant bottom-up and top-down information respectively. The face priming task involved presenting a distractor face and then a target face. The congruence of the prime to the probe face was equally balanced across trials in terms of both the relevant and irrelevant dimensions. There was a negative priming effect (worse performance) when the value of the relevant and/or irrelevant dimension was congruent between the distractor and the target, apart from when the age (irrelevant dimension) of the distractor was congruent with that of the target in the sex *task* – then there was a positive priming effect (better performance).

The presence of priming effects in this experiment was interpreted as evidence that sex and age information from centrally presented but experimentally irrelevant faces is automatically encoded. The presence of interference from the irrelevant dimension indicates that the measured effect cannot occur at the response level – again, because *task* was a between-subjects factor. The negative priming effect of sex in the Age *task,* coupled with the positive priming effect of age in the Sex *task*, was interpreted as indicating that sex information is relevant to age judgements and thus needs to be inhibited, but that age information is not relevant to sex judgements and thus does not need to be inhibited. This suggests that sex information forms a basis for age judgements but not vice versa.

The word priming experiment involved showing participants a word relating to either the task-relevant or task-irrelevant dimension prior to the target face. The prime was a valid cue, in the sense that it was congruent with the probe on 75% of trials. The experiment showed that congruent top-down cues facilitated performance relative to invalid cues – regardless of whether the prime was to the relevant or the irrelevant dimension. This suggests that top-down information can affect sex and age judgments, and provides further evidence for their integrality – since semantic cues about one dimension affect processing of the other.

The result suggests that predictions about an upcoming dimensional value – even if irrelevant to the task – facilitates perceptual decision making through the pre-activation of some sensory and/or representational perceptual units associated with the target response. The experiment suggests that sex and age are integrated at the sensory and/or representational level. I also found tentative evidence to suggest that semantic cues to sex affect performance more than those to age.

Overall, the priming experiments showed that sex and age processing can be influenced by bottom-up and top-down cues from distractor faces and words respectively. The experiments provide further evidence that sex and age are asymmetrically integrated, such that sex information has more of an impact on age discrimination than vice versa. Sex information may therefore be more salient or may be more highly prioritised in perceptual decision making. The results suggest again that asymmetrical interference cannot be accounted for at the level of response selection.

I reported a task-switching experiment in chapter 7 in which the *general* task-switching effect was asymmetrical for each *set* and was modulated by task-difficulty. Switch costs were found for the easier *set* and not the harder *set.* This suggests that the relative difficulty of maintaining and scheduling task-sets is dependent upon stimulus properties – occurring more at the level of sensation than representation. There was a bidirectional, symmetrical *specific* task-switching effect, suggestive of competition at the level of response selection. The symmetry of this cost suggests that, by the stage of response selection, the asymmetry in perceptual integration between sex and age is resolved.

I also found evidence to suggest that age discrimination can be more easily automated than sex discrimination – that age-discrimination performance improves over the course of a block more so than sex-discrimination, in keeping with previous findings (Wiese et al., 2008). This effect seemed to be independent of task-difficulty and the direction of the asymmetry in the general switching cost for each *set.* This set of findings suggests that automation in this task may have more to do with general face-processing mechanisms and strategies than with stimulus properties.

8.3 Stroop findings

The Stroop effect in the current experiments measured Stroop-like compatibility between the values of sex and age. Compatible faces (young females and old males) were processed more efficiently in all five experiments compared to Incompatible faces (young males and old females). A robust Stroop effect was found in each experiment, and a significant Stroop effect was found for each of the three *sets* in the majority of cases. Finding this effect so consistently across five experimental paradigms and three stimulus *sets* strongly confirms this Stroop-like compatibility effect, at least when making forced-choice, binary categorisations on static faces on a screen.

To facilitate the interpretation of this effect, I combined the data across experiments and analysed the Stroop effects as a whole. This was possible because the analysis was the same in each experiment. I took the Stroop data from each experiment and combined it – adding in the factor of *experiment*. This gave me difference scores for each participant in each experiment. I submitted these differences scores to a 2x3x5 between-subjects ANOVA with *task, set* and *experiment* as factors. I did the same for the baseline task-difficulty analyses – using IES as the dependent measure as per the individual experiments.

In both analyses, *task* was treated as a between-subjects measure – despite being a within-subjects measure in the task-switching experiment. This was accounted for by creating a novel identifier for data associated with each *task* for each participant in the task-switching data. This makes the analysis of *task* slightly more conservative in these omnibus analyses – but they were carried out to help illustrate the effects more easily rather than to confirm or disconfirm hypotheses, and so this increase in conservatism was deemed acceptable.

8.3.1 Omnibus tests

8.3.1.1 Omnibus Stroop results

This analysis showed a significant, positive Stroop effect, interpreted as a y-intercept significantly different from zero (F(1,1317) = 499.96, $p < 0.001$, $h^2 = 0.275$, M = 65.7 (2.96)). This effect was modulated by *task* such that there was a larger effect for the Age than the Sex *task* (F(1,1317) = 9.73, p = 0.002, $h^2 = 0.007$, Sex M = 57.15 (4.43), Age M = 74.14 (3.91)), and by *set* (F (2,1317) = 20.38, $p < 0.001$, $h^2 = 0.03$, Facegen M = 90.94 (4.69), MPI M = 58.56 (4.71), Lifespan M = 50.71 (5.09)). FLSD (13.41) revealed that the Stroop effect was significantly larger for the Facegen *set* than both the MPI and Lifespan *set*, which were not significantly different from one another.

These effects were qualified by a significant interaction between *task* and *set* (F(2,1317) = 44.82, $p < 0.001$, $h^2 = 0.064$). Follow-up analyses showed significant differences between *tasks* for all three *sets*. The Stroop effect was significantly larger for the Age than the Sex *task* for both the Facegen (t(371) = 8.85, $p < 0.001$, d = 0.85, Sex M = 54.91 (5.02), Age M = 135.55 (7.61)) and Lifespan *sets* (t(363) = 2.75, p = 0.006, d = 0.26, Sex M = 32.63 (9.2), Age M = 62.67 (5.85)), and significantly larger for the Sex than the Age *task* for the MPI *set* (t(372) = 5.58, $p < 0.001$, d = 0.52, Sex M = 82.63 (8.04), Age M = 30.13 (4.88)) (see figure 8.1).

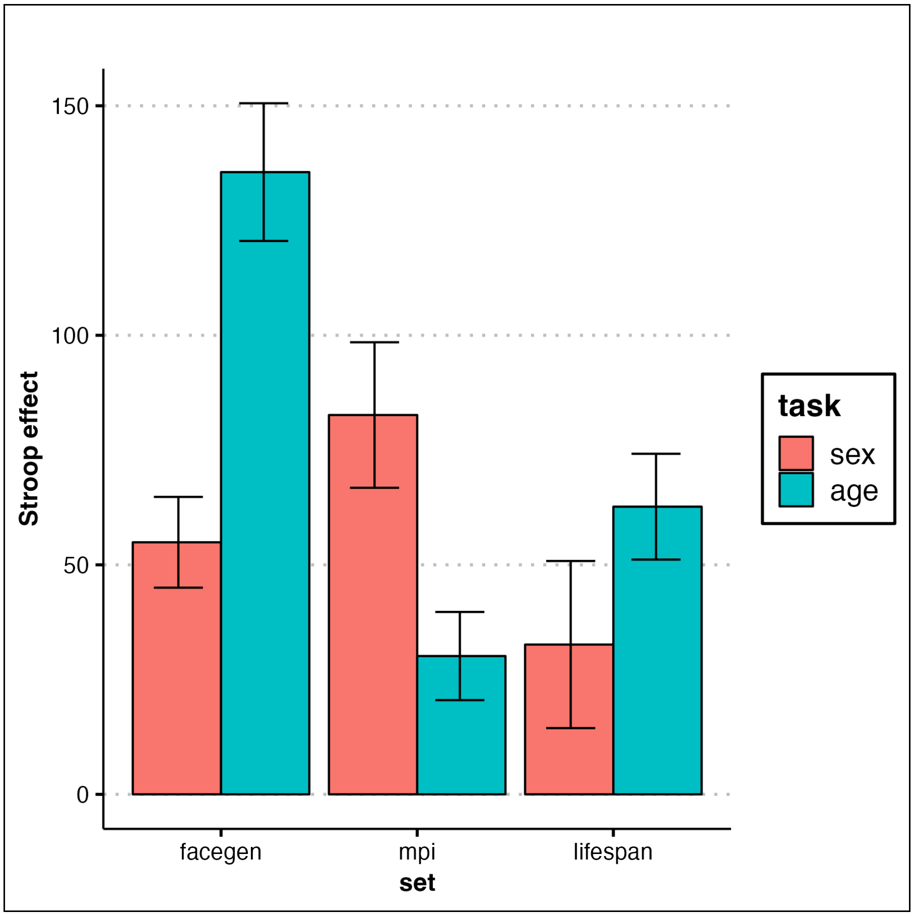


Figure 8.1: The omnibus Stroop results for each *set*, split by *task.* Positive values indicate Stroop effects.

There was no main effect of *experiment* (p = 0.25) and no interactions involving *experiment* (all p > 0.1).

8.3.1.2 Omnibus task-difficulty results

There was a significant main effect of *set* (F(2,1362) = 15.79, p < 0.001, h² = 0.023, Facegen M = 814.26 (9.38), MPI M = 822.3 (9.25), Lifespan M = 873.6 (9.82)), *task* – such that the Sex *task* was harder than the Age *task* (F(1,1362) = 33.03, p < 0.001, h² = 0.024, Sex M = 884.12 (7.9), Age M = 829.87 (7.64)), For the main effect of *set,* FLSD (26.31) showed that the Lifespan *set* was significantly more difficult than both the Facegen and MPI *sets*, and that these two *sets* were not significantly different from one another.

There was a significant interaction between *task* and *set* (F(2,1362) = 55.43, $p < 0.001$, h² = 0.075). Follow-up analyses showed that the Sex *task* was harder than the Age *task* for the Facegen *set* (t(455) = 4.96, $p < 0.001$, d = 0.41), and the Age *task* was harder than the Sex *task* for both the MPI (t(468) = 8.95, $p < 0.001$, d = 0.82) and the Lifespan *sets* (t(455) = 4.96, $p < 0.001$, d = 0.46) (see figure 8.2).

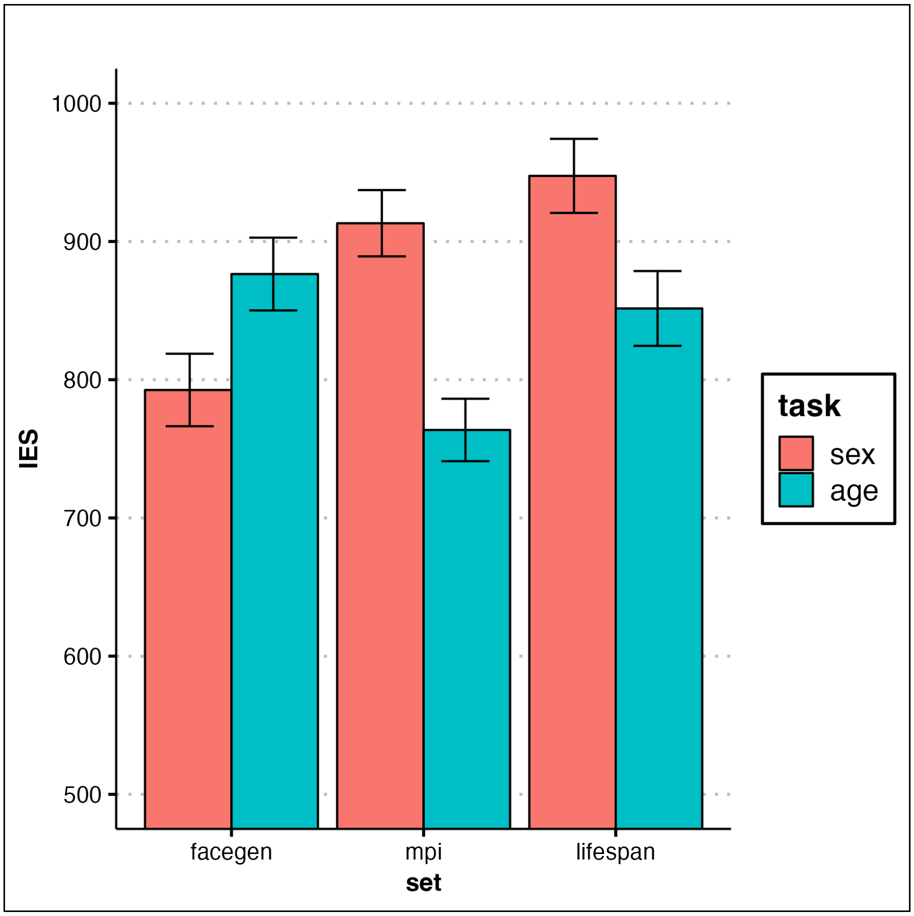


Figure 8.2: IES for each *set*, split by *task* shows interaction between *task* and *set*.

8.3.2 Stroop discussion

Looking at the Stroop effects separately, there were significant differences between *tasks* within *sets* in some experiments, and not in others. In every case, the Stroop effect was larger in the Age than in the Sex *task* for the Facegen *set* and the Lifespan *set*, and was larger in the Sex than in the Age *task* for the MPI *set*. These differences between *tasks* within each *set* were sometimes significant in individual experiments, and were significant in the omnibus *test*.

Compared to some of the main experimental effects, the Stroop effects seemed to be modulated by baseline task-difficulty to a greater extent. Specifically, the larger Stroop effect tended to occur in the more difficult *task* – that is, in the more difficult *task*, there was a greater processing advantage of compatible compared to incompatible faces. The exception to this was for the Lifespan *set*, in which the larger Stroop effect was in the easier *task*.

The Lifespan *set* was also the most difficult *set* – suggesting that the stimulus dimensions were less discriminable for the Lifespan compared to the other two *sets*. It has been noted throughout the current work that the Stroop effect seems to vary with task-difficulty between *sets* – suggesting that the effect emerges at a lower, more sensory level of perceptual processing. The reversal of the Stroop effect between the two photographic *sets*, and the finding that the Lifespan *set* is more difficult than the MPI *set*, may tell us something about the baseline discriminability of the stimulus dimensions between these *sets*.

Taking task-difficulty as a proxy for discriminability, we can argue that the dimensions are less discriminable in the Lifespan *set* than for the other *sets*. If discriminability is low, it suggests that the differences between values of a dimension are less distinguishable from one another. If discriminability of the dimensions was low, it may have been relatively difficult to passively or actively extract cross-dimensional, compatible information to facilitate a relevant judgement. This may mean that the encoding of information from the irrelevant dimension – which may facilitate a task-relevant judgement – may be impaired.

It is worth noting the physical differences between the MPI and Lifespan *sets*, specifically the difference in quality. The MPI faces seem to have been controlled more carefully, with the images captured on a better camera and under better and more uniform lighting conditions (see section 3.4). The Lifespan face database may not be as well suited to social vision research aimed at discriminating dimensions and values under time-pressured conditions – the discriminability between these dimensions and values may not be sufficiently high.

This reasoning – that the discriminability of the dimensions and values is low in the Lifespan *set* – may explain smaller Stroop effects in some experiments, but it does not explain their reversal compared to the MPI *set*, or the fact that the larger Stroop effect is in the harder *task* for the MPI and Facegen *sets*, but the easier *task* for the Lifespan *set*. I cannot explain this reversal.

I have speculated elsewhere (e.g. section 4.5.6) that the larger Stroop effect in the harder *task* (in the Facegen and MPI *sets*) may be explained in terms of the speed-of-processing account (SoPA). According to this account of the phenomenon, Stroop interference occurs for each *task*, but simply has more time to manifest for the harder dimension. When carrying out the harder *task*, the face in each trial is on-screen for longer (because the task is harder) and there is simply more time for the passive or active recruitment of compatible cross-dimensional information for the judgement.

8.3.2.1 Stroop and the shared signal hypothesis

The Stroop results in the current experiment suggest the existence of a “shared signal” between the values of Male and Old on the one hand, and Female and Young on the other. The current Stroop results show asymmetric integrality between sex and age which seems to vary as a function of task-difficulty or dimensional discriminability. This variation suggests that the shared signal between the values of these dimensions originates at the sensory level – that the shared signal can be explained in terms of overlapping physical features (operating on the perceptual system from the bottom-up), rather than overlapping semantic features (operating on the perceptual system from the top-down).

8.3.3 Stroop conclusions

The Stroop effect is a measure of integrality: it shows that sex and age are integrated at the value-level. Its mitigation by task-difficulty may highlight that signals from the task-irrelevant dimension have more scope to interfere when the task is more difficult, and less time when it is easier. Variation in the strength of the Stroop effect may be driven by the discriminability of the objects dimensions and may result from the recruitment of task-irrelevant but perceptually integral information from across dimensions.

Boenke et al. (2009) argued that the Stroop effect occurs earlier than the Garner effect, and pertains to within-presentation level perceptual integration whilst the Garner effect pertains to between-presentation level integration. This seems consistent with the current Stroop results insofar as they appear to be occurring at an early, sensory level.

8.4 Key overall findings

8.4.1 Asymmetric perceptual interference between the dimensions of sex and age

The current results show through converging operations that sex and age are asymmetrically integrated in human face processing. Asymmetric perceptual interference was observed in the Garner effect, face priming and word priming effects, the general task-switching effect and the Stroop effect. Asymmetric interference was not noted in the flanker effect. The latter finding suggests that sex and age information from concurrently presented distractor faces does not contribute to perceptual interference between sex and age in the processing of a target face. However, methodological issues may have mitigated the emergence of this effect (see section 5.5.4)

The former findings suggest that there is asymmetric interactions between sex and age at the dimensional-level (Garner effect) and at the value-level (Stroop effect). Asymmetric effects in target processing emerge as a function of previously encoded irrelevant dimensional information (face priming), and previously encoded semantic information (word priming). Keeping the irrelevant task-set (the perceptual and motoric demands associated with a given task) in working memory also elicits asymmetric interference effects.

For the photographic *sets*, this asymmetry was always such that sex information had a larger, more negative effect on age processing than vice versa. Exceptions to this homogeneity between the photographic *sets* include differences in the Stroop effect (see figure 8.1) and in the direction of the word priming effects (see figure 6.14) – but these two effects show asymmetric facilitation rather than interference. Overall, however, the tendency of sex information to interfere with age processing more than vice versa stands.

For the Facegen *set*, asymmetrical interference was such that sex interfered with age more than vice versa in the Garner effect. Semantic cues to sex also had more of a positive effect on age processing than vice versa for the Facegen *set*. For the other effects, age interfered with sex processing more so than vice versa. The Garner and word priming effects may be indices of interference at a higher, more semantic (less visual) level of processing – more insulated from the particular physical properties characterising each *set*. This would explain why these effects were similar across all three *sets* whilst the other effects differed. This suggests that the asymmetrical relationship between sex and age perception has something to do with the organisation of the face processing system at the semantic level.

8.4.2 Interference at different perceptual levels

The current results suggest that the different interference effects measured occur at different stages of perceptual processing. I argue that the extent to which an effect is mitigated by baseline task-difficulty between *sets* is a measure of the perceptual level at which the effect emerges. Specifically, the Garner effect and the specific task-switching effect seem to occur at higher, more semantic, and less visual levels of representation.

The face priming effect, word priming effect, Stroop effect and general task-switching effect were all mitigated by stimulus *set*, and can therefore be said to be affected by the physical stimulus properties – with interference occurring at lower, less semantic, and more visual levels of processing. The word priming effect – which would be expected to occur at a higher level – may have been modulated by the physical differences between *sets* due to the tuning effect of semantic information on the early visual system.

The flanker effect (Yeh & Eriksen, 1984) and the specific task-switching effects (Kray & Lindenberger, 2002) are supposed to index interference at the response level. I found no evidence of a flanker effect and found specific task-switching effects, but no asymmetry therein. These findings, and the fact that *task* was a between-subjects factor for all but the task-switching experiments, suggests that asymmetric interference processing between sex and age does not occur at the response selection level. Taken together, these findings suggest that asymmetric interference between sex and age occurs at sensory and semantic levels of representation.

Interference appears to arise at the response level (specific task-switching), but any asymmetries may have been resolved by that stage.

8.4.3 Speed-of-processing account explains only some of the results

A speed-of-processing account (SoPA) has previously been put forward to account for asymmetrical perceptual interference between facial dimensions (Graham & LaBar, 2007; Melara & Mounts, 1993). According to this account, easier discrimination judgements are made before information from the harder to discriminate dimension can interfere. Reversing the discriminability of the dimensions (or the difficulty of the experimental tasks) has been shown to reverse the pattern of asymmetrical integrality that occurs between the dimensions (Graham & LaBar, 2007).

The SoPA can generally explain the Stroop effects in the current results (with exceptions, see section 8.3.2), but it cannot explain the Garner effect (which was uniform across *sets* despite differences in task-difficulty), the priming effects (for which visual and semantic cues to either the harder or the easier *task* had the greatest effect on performance across *sets* with the same pattern of task-difficulty) or the general task-switching effect (where there was greater interference of the harder dimension on the easier dimension). Contrary to the SoPA, it was often the case that the harder *task* (Sex, for the photographic *sets*) interfered more with the easier *task* than vice versa.

Within the current work, no attempt was made to reverse the pattern of task-difficulty within *sets*. It is not clear whether doing so would reverse the patterns of observed results. It is only similarities and differences across *sets* in terms of task-difficulty and the corresponding similarities and differences in the observed effects that have led to the questioning of the SoPA. I argue below (section 8.5) that consideration of the reasons *why* the processing of certain dimensions is more or less difficult provides a clearer picture of how asymmetric processing arises.

8.4.5 Sex processing is less efficient than age processing in photographs of faces

For photographs of people, sex processing was less efficient than age processing across both photographic *sets* in all five experiments. This pattern of difficulty was reversed for the artificial Facegen *set*. This finding may suggest that artificial stimuli are processed differently to photographic stimuli, but more research is needed to investigate that hypothesis. The consistency of the finding between the MPI and Lifespan *sets* allows me to argue that sex processing is more difficult (less efficient) quite strongly, particularly since this is consistent with the previous literature (Fitousi, 2020; Quinn & Macrae, 2005).

8.4.6 Evidence of automatic extraction of sex and age information from faces

The primary source of evidence for such information processing comes from the face priming experiment, in which a negative priming effect was found in most conditions, and a positive priming effect for the effect of irrelevant age information on sex judgements. This implies, in Tipper's words "that ignored objects may receive analysis to a categorical level" (Tipper, 1985). The result suggests that the values of sex and age exhibited by distractor faces were processed to the extent that they could interfere with target processing – it suggests that one cannot ignore the age or sex of an irrelevant distractor face, even when that dimension is completely unrelated to the current task.

This finding is bolstered by the post-hoc automation analysis in the task-switching experiment (see figure 7.6), which suggests that sex and age processing increase in efficiency over time, particularly age judgements. This is in line with the findings of Wiese et al. (2008, 2012), who suggested that age processing may proceed in a faster and more automatic fashion than sex processing.

8.4.7 Sex and age perception may differentially rely on top-down and bottom-up processing respectively

In line with the argument that sex and age perception may differentially rely upon configural and featural processing respectively, I tentatively put forwards the idea that the former is more reliant upon top-down, and the latter on bottom-up processing respectively. This hypothesis is bolstered by the facts that, for photographic stimuli, age processing appears to be more easily automated and more efficiently processed – with bottom-up processing being considered to be faster and more automatic than top-down processing (e.g. Theeuwes, 2018).

Similarly to the configural/featural reliance hypothesis, I argue that this difference between sex and age is a matter of processing strategy, and must be considered in relative terms. That is, sex perception is *typically* more reliant upon top-down processing *relative* to age perception, and age perception is *typically* more reliant upon bottom-up processing relative to sex perception. By “more bottom-up”, for example, I mean more reliant upon incoming sensory information about the face, whilst “more top-down” means more reliant upon pre-existing semantic information.

This difference is thought to be an inherent aspect of the fact that configural templates (upon which sex perception is argued to be typically more reliant) are more complex and holistic forms of representation, and can therefore be activated less easily using purely sensory data. It fits into an overarching conception of the sensory-semantic axis of the perceptual processing hierarchy, by which I suggest that if something is “more semantic” it is “less sensory” and vice versa.

I do not state this hypothesis in as strong terms as the configural/featural hypothesis, despite supposing that the former follows on from the latter. However considering asymmetries in perceptual integrality in terms of differential reliance upon sensory and semantic processing may prove fruitful in the future. I focus for the remainder of the discussion upon the more well-evidenced hypothesis relating to differential configural and featural processing

8.5 Interpretation

8.5.1 Current face processing models

The perceptual relationship between sex and age remains under-investigated. It has been suggested that age processing is uncommonly investigated within the perceptual integrality literature because it is not easily categorised as variant or invariant (Dagovtch & Ganel, 2010) – with this distinction characterising the most recent and most neuroanatomically plausible dual-route model of face perception (Haxby et al., 2000, 2002; Duchaine & Yovel, 2015). Neither sex nor age are dealt with in depth by the Bruce and Young (1986) model, and neither are dealt with specifically by the Bartlett model (Bartlett et al., 2003), which does not appear to be as widely discussed as either the Bruce and Young or Haxby models.

In terms of the Bruce and Young model, we might characterise sex and age as being involved in both identity-specific and identity-nonspecific processing. Given this, we should expect them to interact in processing – both being processed along the same cortical routes. The same can be said of the Haxby model, if each are considered invariant dimensions.

One could say that this is because sex is *more* identity-specific, and *more* invariant than age – since age only characterises a person's identity at a given time, changing throughout the lifespan. This characterisation does permit the current results to be interpreted in terms of these models – since it has been argued that identity-specific cues interfere more with identity-nonspecific cues more so than vice versa (Schweinberger & Soukup, 1998), and invariant cues interfere with variant cues more than vice versa (e.g. Karnadewi & Lipp, 2011).

These two models, however, suggest that the processing of the respective types of information with which they deal stems from their being processed according to different perceptual or neuroanatomical systems – that sex and age processing occur within the systems specialised for the processing of identity-specific and identity-nonspecific information (Bruce & Young, 1986), or variant and invariant information (Haxby et al., 2000, 2002). If this were the case, and the above interpretation were correct, we should expect that the asymmetrical effects would not be so significantly modulated by stimulus *set*.

That is, we should expect the effects which characterise the interference processing relationship between sex and age to be conserved across *sets,* regardless of the particular stimuli about which judgements were being made. This was not the case, and some effects differed between *sets* and others did not. This suggests that the relationship of integrality between these two dimensions may be influenced by something else.

I argue that the asymmetrical relationship between these dimensions may be best characterised by their differential reliance upon configural and featural information – in line with the Bartlett model. This explains why there were differences between the *sets*, as the dimensional discriminability of the different *sets* may have led to differences in the relative reliance upon these types of processing in the participants processing strategy.

The distinction between configural and featural processing may be more flexible than that between identity-specific/identity-nonspecific and variant/invariant in information processing. The relative reliance upon configural and featural information may vary more as a function of stimulus properties than the identity specificity or changeability of the dimensions. As such, the distinction permits different dimensions to be processed according to different strategies depending upon the information available. The distinction between configural and featural processing strategies seems more domain-general, putting the type of process before the type of information with which the process is supposed to deal. Proposing that different dimensions are processed according to a differential reliance on configural and featural processing also does not exclude the possibility that the face processing system is set-up along the lines proposed by Bruce and Young or Haxby et al., it merely proposes a mechanism by which such specialised functioning might take place.

Another issue facing the interpretation of the current results within the framework of the dual-route face processing models is my suggestion that different effects emerge at different levels of processing – in particular the findings that top-down information might asymmetrically affect the two dimensions (word priming effect). To account for such phenomena, it is necessary to consider how the face processing system might be open to modulation from pre-existing semantic representations.

Below, I expand upon the hypothesis that asymmetric integrality between sex and age can be thought of in terms of their differential reliance upon configural and featural processing. I refer to configural processing as a “deeper” type of processing than featural processing – which is relatively “shallower”. In referring to the hypothesis, I refer to the depth-of-processing. I attempt to explain the integration of sex and age information at lower sensory, and higher semantic levels referring to the structural reference hypothesis (Ganel et al., 2005) to explain sensory integration, and the parallel-contingent hypothesis (e.g. Johnstone & Downing, 2017) to explain semantic integration. I nest these ideas within what I have described as the depth-of-processing account (DoPA).

8.5.2 The Depth of Processing account (DoPA)

In section 1.5 I reviewed evidence which suggests that sex perception is more reliant upon configural processing, and age perception on featural processing – providing evidence for a neurological distinction between the two types of processing, such that (in broad terms) the former processing may occur over the right hemisphere and the latter over the left hemisphere (Scott & Nelson, 2006). I suggest that sex/configural processing is “deeper” than age/featural processing, that it requires greater computational and attentional resources in order for judgements to be made.

None of the experiments that I have reported in the thesis test this suggestion that sex processing is more reliant on configurational processing and age on featural processing directly, and so I do not claim to have provided any evidence for the idea. I suggest, however, that taking this as an assumption provides a framework within which the current results can be interpreted.

8.5.2.1 Structural reference hypothesis of sensory-level integration

This hypothesis originated in research which found that variations in facial identity interfered with emotional expression judgements but not vice versa in a Garner speeded classification task. This was the case unless the identity was familiar, in which case the interference was bidirectional (Ganel et al., 2004). The authors’ argued that observers have the capacity to use the structure of the face to compute changes that occur within it – that invariant structures provide reference for the computation of variant structures.

A key part of their argument relates to idiosyncrasy. Emotions vary idiosyncratically, but the better a representation an observer has of a given target, the more constrained their knowledge of the range of emotional expressions that the target conveys. It is due to this that, once an identity is known, variations in emotion interfere with identity processing – the observer can use their knowledge of the idiosyncratic repertoire of emotional expressions to constrain their computation of the targets’ identity.

I suggest that sex acts as a structural referent for age based on the assumption that the former relies on configural information and the latter on featural information, and based on the idea that the visual cues to age vary as a function of sex (Kloth et al., 2015), and vary idiosyncratically between individuals. The cues to sex vary between individuals too, but the cues to age vary within sex as well. In this respect, cues to age vary as a function of sex, cues to which vary as a function of identity – age is integrated into sex, sex is integrated into identity.

I argue that the features which characterise age are embedded within the structural configuration which characterises sex. In this, I suggest that sex information *underlies* age information. This is in keeping with the finding of Quinn and Macrae (2005) and Wiese et al. (2008) that age information can be processed without the processing of sex information taking place, but sex information cannot be processed without the processing of age information taking place.

In keeping with the assumption about depth-of-processing, these findings suggest that - when processing sex - one must process the face at a level which includes the processing of age, but not vice versa. This is why I suggest that sex processing involves “deeper” processing than age processing. As such, when processing sex, the depth-of-processing at which the perception takes place incorporates the relatively shallower age processing. When processing age, however, one need not go to that level of depth – processing the face only on the relatively superficial level necessary to extract the featural information necessary to make the judgement.

As discussed above (chapter 4) this can be thought of using the concept of templates (e.g. Yuille, 1991). Our representations of sex and age may be based on templates against which incoming stimulus data are compared. A sufficient match between the bottom-up, physical information and the representational template permits a judgement. Applying this idea to the structural-reference hypothesis, featural templates associated with age judgements are embedded within configural templates associated with sex judgements – with different patterns of featural templates being embedded within the different templates associated with each sex, since males and females age differently.

This line of thinking explains integration at the sensory level – with sensory processing of sex and age being characterised in terms of depth-of-processing. Age judgements can be carried out on the shallow level of matching inputs to featural templates, but these templates are embedded within deeper configural templates (which include sensory-level representations of sex). Irrelevant variations in sex mean irrelevant variations in deep configural templates – this therefore slows down age processing, since the featural processing that needs to be undertaken needs to do so with reference to a different underlying configural template.

Sex judgements are carried out on the deep level of matching inputs to configural templates, over which are superimposed featural templates (which include sensory-level representations of age). Irrelevant variations in age means irrelevant variations in shallow featural templates – this therefore slows down sex processing, since the configural processing that needs to be undertaken is rendered more difficult by relatively superficial changes in featural information. Because this variation occurs on a shallower level, it affects sex processing less than variations on the deeper level affects age processing.

I suggest that the activation of these templates (featural or configural) can occur through bottom-up or top-down activation – which explains why the sensory or word priming of irrelevant information can affect target judgements. Age templates may be more “deformable” (Yuille, 1991) and less “static” (Atkinson et al., 2005) than sex templates. This is because age changes both idiosyncratically and as a function of sex, and because the visual cues which characterise a given persons’ age change continuously throughout the lifespan whereas they tend to remain more static for sex.

In this respect, age is a more variant dimension than sex with respect to a given individual over time. It therefore makes sense that representations of age are embedded within representations of sex at the visual level. I believe that the mechanism by which this embeddedness is instantiated is through the differential reliance of these types of perception on configural and featural processing. Barring a nativist interpretation of the structure of the visual system, it seems likely that the systems’ sensitivity to the embeddedness of age within sex also exists at a higher, less visual, and more semantic level of processing.

8.5.2.2 Parallel-contingent hypothesis of semantic integration

This hypothesis has been used to describe the asymmetric perceptual relationship between cues of body weight and sex, such that variations of the latter affect the processing of the former but not vice versa in a speeded classification task (Johnstone & Downing, 2017). By this hypothesis, it is supposed that the processing of two dimensions occurs in parallel, but that the outputs of the computations of one are passed onto and influence the processing of the other.

Applying the hypothesis to sex and age, the current results suggest that the sex of an individual cannot be ignored when making an age judgement even when it is task-irrelevant because information processing about sex influences parallel processes concerned with age. I argue that this reflects a hierarchical organisation of sex and age at non-visual, semantic levels of organisation in the brain – that knowledge of this fact makes age processing contingent upon sex processing (see figure 6.10).

8.5.2.3 Depth-of-processing

Evidence from the current results which suggests that sex involves deeper processing is that sex was consistently less efficiently processed than age for the photographic stimuli. I also found that age was more easily automated than sex processing in the task-switching experiment. The simplest explanation I can think of to explain these findings is that sex is more complex than age to process. I think that age can be processed based on relatively superficial cues, and that our representations of age are more flexible. This means that sensory input cues can be more easily matched to sensory and/or semantic level representational templates – our representations of age can be deformed more easily to match the incoming information.

I think that because sex is more binary in the “real” visual world than is age – which is binarised in the current experiments but which is continuous in reality – our representations of sex are more fixed. When making a sex judgement, then, one must process the face to a deeper level to make the judgement – whereas the binarisation of age into Young or Old can happen with relatively shallow processing, since our representations of those categories are more flexible.

The relative flexibility of these categories of age is evidenced by the fact that they change across the lifespan and with respect to context. For example, a thirty-year-old may be represented as “old” to a five-year-old but as “young” to a sixty-five year-old. Similarly, a thirty-year-old might be regarded as “young” for a university professor, but “old” to be sucking their thumb. Representations of age vary as a function of the age of the observer in a meaningful way, whereas representations of sex do not vary as a function of the sex of the observer.

This concept of depth, which I use to refer to the axis from featural to configural information processing operations, can also be used to refer to the axis from bottom-up to top-down – or from sensory to semantic processing. “Deeper” processing may involve more reliance upon semantic representations, whilst “shallower” processing may involve less semantic representation and more visual representation. In this, my supposition that sex perception involves “deeper” processing includes the notion that sex perception has more of a semantic component relative to age, which can be more easily perceived using purely sensory cues and requires less semantic information to be processed.

8.6 Conclusions

I argue that age representations are embedded within sex representations at sensory and semantic levels of representation, and that the structure of the hierarchical system which differentiates perception of the two dimensions is organised on the axis of increasing complexity. I argue that sex perception is more complex than age perception, and that the mechanism characterising this difference in complexity relates to the depth of perceptual processing necessary to reach decisions about the respective dimensions. The relative rigidity of visual and semantic representations of sex makes reaching decisions about a person’s sex more difficult than reaching decisions about their age.

This complexity relates to the extent to which configurational and/or semantic processing is necessary. For age perception, configurational processing is not necessary – only featural processing, with that being a relatively simpler, “shallower” type of processing, it also requires less semantic processing and can be carried out using sensory cues more easily than sex processing. This simplicity is what makes it easier to carry out. Sex processing requires configurational and/or semantic processing.

Configurational and featural processing can be carried out in parallel (Amishav & Kimchi, 2010; Kimchi & Amishav, 2010), but featural processing is contingent upon the underlying configurational structure of the object. Therefore, configurational processing interferes with featural processing more so than vice versa. This is why asymmetric interference between these two dimensions occurs.

Although sex and age processing tend to rely differentially upon these processes, each can be carried out using the other process. In this, configurational and featural processing are processing strategies. I argue that the reason why the patterns of task-difficulty (which may be an index of depth-of-processing) are reversed for the Facegen *set* is because the dimorphism between males and females is caricatured – permitting a featural processing based perceptual strategy to be more effective. At the same time, the features which characterise age are less familiar to the participant – since the stimuli are unfamiliar – and so the use of relatively shallow featural processing to make decisions is more difficult.

It seems that perceptual decisions relating to these decisions can be modulated by bottom-up and top-down signals. This may occur through the activation of representations of these dimensions and their respective values at more semantic and/or more visual levels of the system. The activation of semantic representations may tune visual receptors towards target values from the top-down (see Henderson et al. 2023). The asymmetric integration of sex and age information seems, however, to occur at sensory and semantic levels – suggesting an asymmetrical hierarchical organisation within which sex information takes priority.

Although much of this is speculative, the ideas provide scope for future testing. For example, I have suggested that featural information may be contingent upon configural information, and that the former may be a structural referent for the latter. Independently manipulating these aspects of a stimulus structure in a speeded classification task may be a test of that hypothesis. Irrelevant variations in configural information should interfere with featural judgements more so than vice versa.

Similarly, I have furthered the hypothesis that the source of perceptual integration may be in the depth-of-processing necessary for the computation of each stimulus dimension – with that being characterised by a differential reliance upon configural and featural information. Other facial dimensions might be characterised in these terms and then subjected to tests of perceptual integrality. If my hypothesis holds, there should be asymmetrical interference such that more configuration-based dimensions such as sex and identity should interfere with more featurally-based dimensions such as age, eye-gaze and emotional expression.

Indeed, such patterns of integrality have already been demonstrated (e.g. Atkinson et al., 2005; Schweinberger & Soukup, 1998), but these have been interpreted in terms of the Haxby model, and invariant against variant dimensions. I propose that the understanding of the face processing system may be based more accurately upon the Bartlett model, and the distinction between configural and featural processing – and that adopting this framework may lead to more consistent findings.

In terms of the extant models of face processing, the current results seem to be best explained by the Bartlett model (Bartlett et al., 2003) and the Freeman and Ambady model (2011). I suggest that the structure of the face processing system is like that proposed by Freeman and Ambady – a multi-levelled hierarchical system within which bottom-up physical signals and top-down semantic signals dynamically converge over time into stable perceptual states through ongoing cycles of mutually excitatory and inhibitory feedback. I add, however, that one of the mechanisms by which this is achieved is through depth-of-processing, that the extraction of sensory information about different stimulus dimensions differentially relies upon featural and configural information.

References

Adams Jr, R. B., Albohn, D. N., & Kveraga, K. (2017). Social vision: Applying a social-functional approach to face and expression perception. *Current directions in psychological science*, *26*(3), 243-248.

Adams Jr, R. B., Albohn, D. N., Hedgecoth, N., Garrido, C. O., & Adams, K. D. (2022). Angry white faces: A contradiction of racial stereotypes and emotion-resembling appearance. *Affective science*, *3*(1), 46-61.

Adams Jr, R. B., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological science*, *14*(6), 644-647.

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron*, *21*(2), 373-383.

Algom, D., Chajut, E. & Lev, S. (2004). A Rational Look at the Emotional Stroop Phenomenon: A Generic Slowdown, Not a Stroop Effect. *Journal of Experimental Psychology: General*, *133*(3), 323–338. https://doi.org/10.1037/0096-3445.133.3.323

Algom, D. & Fitousi, D. (2016). Half a Century of Research on Garner Interference and the Separability–Integrality Distinction. *Psychological Bulletin*, *142*(12), 1352–1383. https://doi.org/10.1037/bul0000072

Allport, G. W. (1954). The nature of prejudice. Cambridge, MA: Addison Wesley.

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance 15: Conscious and nonconscious information processing* (pp. 421–452). The MIT Press

Amishav, R. & Kimchi, R. (2010). Perceptual integrality of componential and configural information in faces. *Psychonomic Bulletin & Review*, *17*(5), 743–748. https://doi.org/10.3758/pbr.17.5.743

Andersen, B. P., Miller, M., & Vervaeke, J. (2022). Predictive processing and relevance realization: exploring convergent solutions to the frame problem. *Phenomenology and the Cognitive Sciences*, 1-22.

Anderson, E., Siegel, E. H., Bliss-Moreau, E., & Barrett, L. F. (2011). The visual impact of gossip. *Science*, *332*(6036), 1446-1448.

Anderson, P. (1957). *Call Me Joe* (pp. 1-33). Atlas Publishing & Distributing.

Aksentijevic, A. & Gibson, K. Psychological complexity and the cost of information processing. *Theor Psychol* **22**, 572–590 (2012).

Albonico, A., & Barton, J. (2019). Progress in perceptual research: the case of prosopagnosia. *F1000Research*, *8*.

Ashby, F. G. & Townsend, J. T. (1986). Varieties of Perceptual Independence. *Psychological Review*, *2*. https://doi.org/10.1037/0033-295x.93.2.154

Ashley, V., & Swick, D. (2019). Angry and fearful face conflict effects in post-traumatic stress disorder. *Frontiers in psychology*, *10*, 136.

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

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in cognitive sciences*, *16*(8), 437-443.

Axelrod, V., & Yovel, G. (2013). The challenge of localizing the anterior temporal face area: a possible solution. *Neuroimage*, *81*, 371-380.

Baddeley, A. D., & Hitch, G. J. (2007). Working memory: Past, present… and future. *The cognitive neuroscience of working memory*, 1-20.

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087-9092.

Baluch, F., & Itti, L. (2011). Mechanisms of top-down attention. *Trends in neurosciences*, *34*(4), 210-224.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological science*, *7*(3), 136-141.

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

Barlow, H. B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology?. *Perception*, *1*(4), 371-394.

Bartlett, F. C. (1949). Social Influences Upon Memory. In W. Dennis (Ed.), *Readings in general psychology* (pp. 277–281). Prentice-Hall, Inc

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive psychology*, *25*(3), 281-316.

Bartlett, J.C, Searcy, J.H & Abdi, H. (2003). What Are the Routes to Face Recognition?. *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes*, 21.

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

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cognition & emotion*, *28*(3), 416-432.

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), 179–190. https://doi.org/10.1037/0022-3514.92.2.179

Behrmann, M., Richler, J. J., Avidan, G., & Kimchi, R. (2014). Holistic face perception.

Beloff, J. (1975). Koestler's philosophy of mind. *Astride the two cultures: Arthur Koestler at*, *70*, 69-83.

Benson, N. C., Butt, O. H., Datta, R., Radoeva, P. D., Brainard, D. H., & Aguirre, G. K. (2012). The retinotopic organization of striate cortex is well predicted by surface topology. *Current Biology*, *22*(21), 2081-2085.

Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: the impact of age-related craniofacial changes on social perception. *Psychological bulletin*, *100*(1), 3.

Biederman, I. (1972). Human performance in contingent information-processing tasks. *Journal of Experimental Psychology*, *93*(2), 219.

Bindemann, M., Burton, A. M., & Jenkins, R. (2005). Capacity limits for face processing. *Cognition*, *98*(2), 177-197.

Boenke, L. T., Ohl, F. W., Nikolaev, A. R., Lachmann, T. & Leeuwen, C. van. (2009). Different time courses of Stroop and Garner effects in perception — An Event-Related Potentials Study. *NeuroImage*, *45*(4), 1272–1288. https://doi.org/10.1016/j.neuroimage.2009.01.019

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *Quarterly Journal of Experimental Psychology*, *66*(12), 2426-2442.

Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., ... & Rosen, B. R. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron*, *17*(5), 875-887.

Brewer, W. F. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 21–90). Cambridge University Press.

Broadbent, D. E. (1958). *Perception and communication*. New York: Oxford University Press.

Brody, L. R., & Hall, J. A. (2008). Gender and emotion in context. *Handbook of emotions*, *3*, 395-408.

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological review*, *64*(3), 205.

Broadbent, D. E. (1977). The hidden preattentive processes. *American Psychologist*, *32*(2), 109.

Brown, E. & Perrett, D. I. (1992). What Gives a Face its Gender? *Perception*, *22*(7), 829–840. https://doi.org/10.1068/p220829

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., Dench, N., & Burton, M. (1993). Effects of distinctiveness, repetition and semantic priming on the recognition of face familiarity. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *47*(1), 38.

Bruce, V. & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Bruner, J. (1992). Another look at new look 1. *American Psychologist*, *47*(6), 780.

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

Brunelli, R., & Poggio, T. (1993). Face recognition: Features versus templates. *IEEE transactions on pattern analysis and machine intelligence*, *15*(10), 1042-1052.

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)?. *Psychologica Belgica*, *51*(1), 5-13.

Buckland, M. K. (1991). Information as thing. *Journal of the American Society for information science*, *42*(5), 351-360.

Buckner, R. L., & Krienen, F. M. (2013). The evolution of distributed association networks in the human brain. *Trends in cognitive sciences*, *17*(12), 648-665.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends in cognitive sciences*, *10*(4), 159-166.

Bukowski, M., de Lemus, S., Marzecová, A., Lupiáñez, J., & Gocłowska, M. A. (2019). Different faces of (un) controllability: Control restoration modulates the efficiency of task switching. *Motivation and Emotion*, *43*, 12-34.

Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and colour information. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *259*(1355), 137-143.

Burton, A. M., & Bruce, V. (1993). Naming faces and naming names: Exploring an interactive activation model of person recognition. *Memory*, *1*(4), 457-480.

Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J. M. (2015). Arguments against a configural processing account of familiar face recognition. *Perspectives on Psychological Science*, *10*(4), 482-496.

Calvert, G. A. (1997). *Functional magnetic resonance imaging studies of lip-reading and audio-visual speech perception* (Doctoral dissertation, University of Oxford).

Capurro, R., & Hjørland, B. (2003). The concept of information. *Annual Review of Information Science and Technology* 37(8), 343-411

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision research*, *51*(13), 1484-1525.

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in human behavior*, *29*(6), 2156-2160.

Casper, C., Rothermund, K., & Wentura, D. (2011). The activation of specific facets of age stereotypes depends on individuating information. *Social Cognition*, *29*(4), 393-414.

Castro, S., Bukowski, M., Lupiáñez, J., & Wodniecka, Z. (2021). Fast or accurate? The change of goals modulates the efficiency of executive control. *Polish Psychological Bulletin*, *52*(1), 49-66.

Cavanagh, P., Caplovitz, G. P., Lytchenko, T. K., Maechler, M. R., Tse, P. U., & Sheinberg, D. L. (2023). The Architecture of Object-Based Attention. *Psychonomic Bulletin & Review*, 1-25.

Chelazzi, L., Libera, C. D., Sani, I. & Santandrea, E. (2011). Neural basis of visual selective attention. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(4), 392–407. https://doi.org/10.1002/wcs.117

Chen, E., Xia, B., Lian, Y., Zhang, Q., Yang, X., & Wang, H. (2022). The role of discriminability in face perception: Interference processing of expression, gender, and gaze. *Attention, Perception, & Psychophysics*, *84*(7), 2281-2292.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295x.82.6.407

Bindemann, M., Burton, A. M., & Jenkins, R. (2005). Capacity limits for face processing. *Cognition*, *98*(2), 177-197.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on information theory*, *2*(3), 113-124.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual review of psychology*, *62*, 73-101.

Chronicle, E. P., Chan, M. Y., Hawkings, 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.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, *36*(3), 181-204.

Clark, V. P., Keil, K., Maisog, J. M., Courtney, S. L. G. U., Ungerleider, L. G., & Haxby, J. V. (1996). Functional magnetic resonance imaging of human visual cortex during face matching: a comparison with positron emission tomography. *Neuroimage*, *4*(1), 1-15.

Clarke, D. D., & Sokoloff, L. (1994). Circulation and energy metabolism of the brain. Basic neurochemistry: molecular, cellular and medical aspects. *ed*, *6*.

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.

Cohen, G. (2000). Hierarchical models in cognition: Do they have psychological reality?. *European Journal of cognitive psychology*, *12*(1), 1-36.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*(6), 671-684.

Crick, F. (1990). *The astonishing hypothesis: The scientific search for the soul.* Touchstone Books

Dagovitch, Y., & Ganel, T. (2010). Effects of Facial Identity on Age Judgments. *Experimental psychology*.

Dalrymple-Alford, E. C., & Budayr, B. (1966). Examination of some aspects of the Stroop color-word test. *Perceptual and motor skills*, *23*(3_suppl), 1211-1214.

Dawkins, R. (1996). *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. WW Norton & Company.

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. https://doi.org/10.1080/13506285.2012.744788

De Haan, M., & Nelson, C. A. (1999). Brain activity differentiates face and object processing in 6-month-old infants. *Developmental psychology*, *35*(4), 1113.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience*, *18*(1), 193-222.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological review*, *70*(1), 80.

Diamond, R. & Carey, S. (1986). Why Faces Are and Are Not Special: An Effect of Expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117. https://doi.org/10.1037/0096-3445.115.2.107

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.

Dretske, F. I. (1981). *Knowledge and the Flow of Information*. The MIT Press

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*(1), 53–78. https://doi.org/10.1348/000712601162103

Duchaine, B. & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, *1*(1), 393–416. https://doi.org/10.1146/annurev-vision-082114-035518

Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, *6*(5), 178-190.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, *87*(3), 272–300. https://doi.org/10.1037/0033-295x.87.3.272

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of experimental psychology: General*, *113*(4), 501.

Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *Plos one*, *15*(4), e0231968.

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES - A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42*(1), 351-362. doi:10.3758/BRM.42.1.351.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual review of psychology*, *48*(1), 269-297.

Eimer, M. (2000). Effects of face inversion on the structural encoding and recognition of faces: Evidence from event-related brain potentials. *Cognitive Brain Research*, *10*(1-2), 145-158.

Ellefson, M. R., Shapiro, L. R., & Chater, N. (2006). Asymmetrical switch costs in children. *Cognitive Development*, *21*(2), 108-130.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding?. *Neuron*, *23*(1), 115-125.

Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/bf03203267

Evans, K. K., Horowitz, T. S., Howe, P., Pedersini, R., Reijnen, E., Pinto, Y., Kuzmova, Y. & Wolfe, J. M. (2011). Visual attention. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 503–514. https://doi.org/10.1002/wcs.127

Eysenck, M.W., Keane, M.T (2015) *Cognitive psychology*. Taylor & Francis

Farah, M. J. (1991). Cognitive neuropsychology: Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology*, *8*(1), 1-19.

Farah, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain specificity in visual object recognition. *Current Directions in Psychological Science*, *1*(5), 164-169.

Farah, M. J., Wilson, K. D., Drain, M. & Tanaka, J. N. (1998). What Is “Special” About Face Perception? *Psychological Review*, *105*(3), 482–498. https://doi.org/10.1037/0033-295x.105.3.482

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of personality and social psychology*, *50*(2), 229.

Fenske, M. J., & Eastwood, J. D. (2003). Modulation of focused attention by faces expressing emotion: evidence from flanker tasks. *Emotion*, *3*(4), 327.

Fifić, M. & Townsend, J. T. (2010). Information-Processing Alternatives to Holistic Perception: Identifying the Mechanisms of Secondary-Level Holism Within a Categorization Paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1290–1313. https://doi.org/10.1037/a0020123

Finger, S. (2001). *Origins of neuroscience: a history of explorations into brain function*. Oxford University Press.

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.

Fitousi, D. (2016). Simon and Garner effects with color and location: Evidence for two independent routes by which irrelevant location influences performance. *Attention, Perception, & Psychophysics*, *78*, 2433-2455.

Fitousi, D. (2020). Evaluating the independence of age, sex, and race in judgment of faces. *Cognition*, *202*, 104333. https://doi.org/10.1016/j.cognition.2020.104333

Freeman, J. B. & Ambady, N. (2011). A Dynamic Interactive Theory of Person Construal. *Psychological Review*, *118*(2), 247–279. https://doi.org/10.1037/a0022327

Freeman, J. B., Stolier, R. M. & Brooks, J. A. (2019). Dynamic interactive theory as a domain-general account of social perception. *Advances in Experimental Social Psychology*, *61*, 237–287. https://doi.org/10.1016/bs.aesp.2019.09.005

Freiwald, W. A. (2020). The neural mechanisms of face processing: cells, areas, networks, and models. *Current Opinion in Neurobiology*, *60*, 184–191. https://doi.org/10.1016/j.conb.2019.12.007

Friedenberg, J. (2009). *Dynamical psychology: Complexity, self-organization and mind*. Isce Publishing.

Frings, C., Wentura, D., & Wühr, P. (2012). On the fate of distractor representations. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 570.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*(9), 1325-1352.

Friston, K. (2012). The history of the future of the Bayesian brain. *NeuroImage*, *62*(2), 1230-1233.

Furth, H. G. (1968). Piaget's theory of knowledge: The nature of representation and interiorization. *Psychological Review*, *75*(2), 143.

Gandolfo, M. & Downing, P. E. (2023). Asymmetric visual representation of sex from facial appearance. *Psychonomic Bulletin & Review*, *30*(2), 585–595. https://doi.org/10.3758/s13423-022-02199-6

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), 854–867.

Ganel, T., Goshen-Gottstein, Y. & Ganel, T. (2004). Effects of Familiarity on the Perceptual Integrality of the Identity and Expression of Faces: The Parallel-Route Hypothesis Revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(3), 583–597.

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. https://doi.org/10.1016/j.visres.2004.06.025

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.

Garner, W. R. (1974). *The processing of information and structure.* Lawrence Erlbaum.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts.* Wiley.

Garner, W. R., Hake, H. W. & Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological Review*, *63*(3), 149–159. https://doi.org/10.1037/h0042992

Garner, W. R., & Morton, J. (1969). Perceptual independence: definitions, models, and experimental paradigms. *Psychological Bulletin*, *72*(4), 233.

Gal, P. M. L. & Bruce, V. (2002). Evaluating the independence of sex and expression in judgments of faces. *Perception & Psychophysics*, *64*(2), 230–243. https://doi.org/10.3758/bf03195789

Gauthier, I., & Logothetis, N. K. (2000). Is face recognition not so unique after all?. *Cognitive Neuropsychology*, *17*(1-3), 125-142.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 431.

George, P. A., & Hole, G. J. (1995). Factors influencing the accuracy of age estimates of unfamiliar faces. *Perception*, *24*(9), 1059-1073.

George, P. A. & Hole, G. J. (1996). The Influence of Feature-Based Information in the Age Processing of Unfamiliar Faces. *Perception*, *27*(3), 295–312. https://doi.org/10.1068/p270295

George, P. A., & Hole, G. J. (1998). The influence of feature-based information in the age processing of unfamiliar faces. *Perception*, *27*(3), 295-312.

George, P. A., & Hole, G. J. (2000). The role of spatial and surface cues in the age-processing of unfamiliar faces. *Visual Cognition*, *7*(4), 485-509.

Gibson, J.J. (1979) *The ecological approach to visual perception.* Lawrence Erlbaum Associates

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and social Psychology*, *60*(4), 509.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350-363.

Gilbert, S. J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive psychology*, *44*(3), 297-337.

Gilead, M., Trope, Y., & Liberman, N. (2020). Above and beyond the concrete: The diverse representational substrates of the predictive brain. *Behavioral and Brain Sciences*, *43*, e121.

Gleick, J. (2011). *The information: A history, a theory, a flood*. Vintage.

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

González-Álvarez, J., & Sos-Peña, R. (2022). Sex perception from facial structure: Categorization with and without skin texture and color. *Vision Research*, *201*, 108127.

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

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), 1198–1214.

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

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1, pp. 1969-2012). New York: Wiley.

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, *15*(8), 536-548.

Grill-Spector, K., Weiner, K. S., Gomez, J., Stigliani, A. & Natu, V. S. (2018). The functional neuroanatomy of face perception: from brain measurements to deep neural networks. *Interface Focus*, *8*(4), 20180013. https://doi.org/10.1098/rsfs.2018.0013

Grill-Spector, K., Weiner, K. S., Kay, K. & Gomez, J. (2016). The Functional Neuroanatomy of Human Face Perception. *Annual Review of Vision Science*, *3*(1), 1–30. https://doi.org/10.1146/annurev-vision-102016-061214

Grossman, E. D., Battelli, L., & Pascual-Leone, A. (2005). Repetitive TMS over posterior STS disrupts perception of biological motion. *Vision research*, *45*(22), 2847-2853.

Harnad, S. (Ed.). (1987). *Categorical perception: The groundwork of cognition.* Cambridge University Press.

Hautekiet, C., Verschooren, S., Langerock, N., & Vergauwe, E. (2023). Attentional switching between perception and memory: Examining asymmetrical switch costs. *Attention, Perception, & Psychophysics*, 1-11.

Haxby, J. V., Hoffman, E. A., Gobbini, M. I., 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. https://doi.org/10.1016/s1364-6613(00)01482-0

Haxby, J. V., Hoffman, E. A. & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*(1), 59–67. https://doi.org/10.1016/s0006-3223(01)01330-0

Hedgecoth, N., Strand, N., & Adams Jr, R. B. (2023). The Intersection of Race, Gender/Sex, and Age in Emotion Perception from Faces and Bodies. *Emotion Communication by the Aging Face and Body: A Multidisciplinary View*, 106.

Helmholtz, H. v. (2009). *Treatise on physiological optics* (1867). In B. F. Gentile & B. O. Miller, *Foundations of psychological thought: A history of psychology* (pp. 136–160). Sage Publications, Inc.

Henderson, M. M., Tarr, M. J., & Wehbe, L. (2023). Low-level tuning biases in higher visual cortex reflect the semantic informativeness of visual features. *Journal of Vision*, *23*(4), 8-8.

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

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

Hole, G., & George, P. (2011). Evidence for holistic processing of facial age. *Visual Cognition*, *19*(5), 585-615.

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature neuroscience*, *3*(3), 284-291.

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.

Hübner, R., Steinhauser, M., & Lehle, C. (2010). A dual-stage two-phase model of selective attention. *Psychological review*, *117*(3), 759.

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

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. NeuroImage, 15, 353–372.

Ito, M., & Komatsu, H. (2004). Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *Journal of Neuroscience*, *24*(13), 3313-3324.

Ito, T. A. & Urland, G. R. (2003). Race and Gender on the Brain: Electrocortical Measures of Attention to the Race and Gender of Multiply Categorizable Individuals. *Journal of Personality and Social Psychology*, *85*(4), 616–626. https://doi.org/10.1037/0022-3514.85.4.616

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature reviews neuroscience*, *2*(3), 194-203.

James, W. (1890). *The principles of psychology, Vol. 1.* Henry Holt and Co. https://doi.org/10.1037/10538-000

Jenkins, R., Lavie, N., & Driver, J. (2003). Ignoring famous faces: Category-specific dilution of distractor interference. *Perception & Psychophysics*, *65*(2), 298-309.

Jersild, A. T. (1927). *Mental set and shift* (No. 89). Columbia university.

Johnson, K. L., & Adams Jr, R. B. (2013). Social vision: An introduction. *Social Cognition*, *31*(6), 633-635.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1-2), 1-19.

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: how covarying phenotypes and stereotypes bias sex categorization. *Journal of personality and social psychology*, *102*(1), 116.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.

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), 1–17. https://doi.org/10.1080/13506285.2017.1334733

Jones, C. R. & Fazio, R. H. (2010). Person Categorization and Automatic Racial Stereotyping Effects on Weapon Identification. *Personality and Social Psychology Bulletin*, *36*(8), 1073–1085. https://doi.org/10.1177/0146167210375817

Jones, G., & Smith, P. K. (1984). The eyes have it: Young children's discrimination of age in masked and unmasked facial photographs. *Journal of Experimental Child Psychology*, *38*(2), 328-337.

Kaas, J. H. (2013). The evolution of brains from early mammals to humans. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(1), 33-45.

Kant, I. (1908). Critique of pure reason. 1781. *Modern Classical Philosophers, Cambridge, MA: Houghton Mifflin*, 370-456.

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.

Kamiyama, A., Fujita, K., & Kashimori, Y. (2016). A neural mechanism of dynamic gating of task-relevant information by top-down influence in primary visual cortex. *BioSystems*, *150*, 138-148.

Karnadewi, F. & Lipp, O. V. (2011). The Processing of Invariant and Variant Face Cues in the Garner Paradigm. *Emotion*, *11*(3), 563–571. https://doi.org/10.1037/a0021333

Katsuki, F., & Constantinidis, C. (2014). Bottom-up and top-down attention: different processes and overlapping neural systems. *The Neuroscientist*, *20*(5), 509-521.

Kay, K. N., & Yeatman, J. D. (2017). Bottom-up and top-down computations in word-and face-selective cortex. *elife*, *6*, e22341.

Kemp, R., Pike, G., White, P., & Musselman, A. (1996). Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception*, *25*(1), 37-52.

Kiefer, M. (2008). Top-down modulation of unconscious “automatic” processes: A gating framework. *Advances in Cognitive Psychology*, *3*(1–2), 289–306. https://doi.org/10.2478/v10053-008-0032-2

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological bulletin*, *136*(5), 849.

Kietzmann, T. C., Swisher, J. D., König, P., & Tong, F. (2012). Prevalence of selectivity for mirror-symmetric views of faces in the ventral and dorsal visual pathways. *Journal of Neuroscience*, *32*(34), 11763-11772.

Kinchla, R. A. (1992). Attention. *Annual Review of Psychology, 43,* 711–742. https://doi.org/10.1146/annurev.ps.43.020192.003431

Kihlstrom, J. F. (1987). The cognitive unconscious. *Science*, *237*(4821), 1445-1452.

Kimchi, R. & Amishav, R. (2010). Faces as perceptual wholes: The interplay between component and configural properties in face processing. *Visual Cognition*, *18*(7), 1034–1062. https://doi.org/10.1080/13506281003619986

Kinzler, K. D., Shutts, K., & Correll, J. (2010). Priorities in social categories. *European Journal of Social Psychology*, *40*(4), 581-592.

Kloth, N., Damm, M., Schweinberger, S. R., & Wiese, H. (2015). Aging affects sex categorization of male and female faces in opposite ways. *Acta psychologica*, *158*, 78-86.

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking—An integrative review of dual-task and task-switching research. *Psychological bulletin*, *144*(6), 557.

Koestler, A. (1964). *The act of creation*. Hutchinson

Koestler, A. (1968). *The ghost in the machine*. Picador

Konorski, J. (1967). Some new ideas concerning the physiological mechanisms of perception. *Acta Biol Exp (Warsz)*, *27*(2), 147-61.

Kossowska, M., Bukowski, M., & Czarnek, G. (2014). Two routes to closure: Time pressure and goal activation effects on executive control. *Polish Psychological Bulletin*, *45*(3).

Kramer, R. S. 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), 115–129. https://doi.org/10.1037/rev0000048

Kray, J. & Lindenberger, U. (2000). Adult Age Differences in Task Switching. *Psychology and Aging*, *15*(1), 126–147. https://doi.org/10.1037/0882-7974.15.1.126

Kray, J., Li, K. Z. H. & Lindenberger, U. (2002). Age-Related Changes in Task-Switching Components: The Role of Task Uncertainty. *Brain and Cognition*, *49*(3), 363–381. https://doi.org/10.1006/brcg.2001.1505

Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: expression, experience, and physiology. *Journal of personality and social psychology*, *74*(3), 686.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press

Langton, S. R., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, *107*(1), 330-342.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human perception and performance*, *21*(3), 451.

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychological science*, *14*(5), 510-515.

Lavie, N. & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197. https://doi.org/10.3758/bf03213897

Lee, T. S. (2003). Computations in the early visual cortex. *Journal of Physiology-Paris*, *97*(2–3), 121–139. https://doi.org/10.1016/j.jphysparis.2003.09.015

Leeuwen, C. V., & Bakker, L. (1995). Stroop can occur without Garner interference: Strategic and mandatory influences in multidimensional stimuli. *Perception & Psychophysics*, *57*, 379-392.

Lemm, K. M., Dabady, M., & Banaji, M. R. (2005). Gender picture priming: It works with denotative and connotative primes. *Social cognition*, *23*(3), 218-241.

Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral cortex*, *11*(4), 287-297.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in cognitive sciences*, *6*(1), 9-16.

Lichtenstein-Vidne, L., Henik, A., & Safadi, Z. (2012). Task relevance modulates processing of distracting emotional stimuli. *Cognition & Emotion*, *26*(1), 42-52.

Liu, J., Harris, A., & Kanwisher, N. (2002). The M100: Face categorization begins within 100 ms of stimulus presentation. *Journal of Vision*, *2*(7), 611-611.

Lockhead, G. R. (1966). Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, *72*(1), 95.

Loewenstein, W. R. (1999). *The touchstone of life: molecular information, cell communication, and the foundations of life*. Oxford University Press.

Luria, A. R. (1970). The functional organization of the brain. *Scientific American*, *222*(3), 66-79.

Luria, A. R. (1932). *The nature of human conflicts.* Liveright.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Macrae, C. N., & Bodenhausen, G. V. (2001). Social cognition: Categorical person perception. *British journal of psychology*, *92*(1), 239-255.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Calvini, G. (1999). Seeing more than we can know: Visual attention and category activation. *Journal of Experimental Social Psychology*, *35*(6), 590-602.

Macrae, C. N., Stangor, C., & Hewstone, M. (Eds.). (1996). *Stereotypes and stereotyping*. Guilford Press.

Maddox, W. T. (1992). Perceptual and decisional separability. In *Multidimensional models of perception and cognition* (pp. 147-180). Psychology Press.

Markman, A. B., & Dietrich, E. (2000). In defense of representation. *Cognitive Psychology*, *40*(2), 138-171.

Marí-Beffa, P., Hayes, A. E., Machado, L., & Hindle, J. V. (2005). Lack of inhibition in Parkinson's disease: evidence from a lexical decision task. *Neuropsychologia*, *43*(4), 638-646.

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

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., Swainson, R., Slessor, G., Hutchison, J., Marosi, D., & Cunningham, S. J. (2015). The simultaneous extraction of multiple social categories from unfamiliar faces. *Journal of Experimental Social Psychology*, *60*, 51-58.

Maurer, D., Grand, R. L. & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260. https://doi.org/10.1016/s1364-6613(02)01903-4

May, C. P., Kane, M. J., & Hasher, L. (1995). Determinants of negative priming. *Psychological bulletin*, *118*(1), 35.

Marzecová, A., Bukowski, M., Correa, Á., Boros, M., Lupiáñez, J., & Wodniecka, Z. (2013). Tracing the bilingual advantage in cognitive control: The role of flexibility in temporal preparation and category switching. *Journal of Cognitive Psychology*, *25*(5), 586-604.

McCarthy, J., Minsky, M. L., & Rochester, N. (1955). A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE.

McClelland, J. L. (1995). Constructive memory and memory distortions: A parallel-distributed processing approach.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, *88*(5), 375.

McClelland, J. L., Rumelhart, D. E., & Hinton, G. E. (1986). The appeal of parallel distributed processing. *MIT Press, Cambridge MA*, *3*, 44.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces?. *Trends in cognitive sciences*, *11*(1), 8-15.

McNamara, T. P., & Holbrook, J. B. (2003). Semantic memory and priming. *Handbook of psychology: Experimental psychology*, *4*, 447-474.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive psychology*, *41*(3), 211-253.

Melara, R. D., & Mounts, J. R. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, *21*(5), 627-645.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of memory and language*, *40*(1), 25-40.

Miller, G. (1956). Human memory and the storage of information. *IRE Transactions on Information Theory*, *2*(3), 129-137.

Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in cognitive sciences*, *7*(3), 141-144.

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior research methods, instruments, & computers*, *36*, 630-633.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/s1364-6613(03)00028-7

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of verbal learning and verbal behavior*, *16*(5), 519-533.

Moser, J. S., Huppert, J. D., Duval, E. & Simons, R. F. (2008). Face processing biases in social anxiety: An electrophysiological study. *Biological Psychology*, *78*(1), 93–103. https://doi.org/10.1016/j.biopsycho.2008.01.005

Mouchetant-Rostaing, Y., & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *Journal of cognitive neuroscience*, *15*(6), 900–910. https://doi.org/10.1162/089892903322370816

Murray, J. E., Machado, L., & Knight, B. (2011). Race and gender of faces can be ignored. *Psychological Research*, *75*, 324-333.

Needham, J. (1932). Thoughts on the Problem of biological Organisation. Scientia 26 (52):84.

Neill, W. T. (1977). Inhibitory and facilitatory processes in selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(3), 444.

Neisser, U. (2014). *Cognitive psychology: Classic edition*. Psychology press.

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

Neumann, O. (1984). Automatic processing: A review of recent findings and a plea for an old theory. *Cognition and motor processes*, 255-293.

Neumann, M. F., & Schweinberger, S. R. (2009). N250r ERP repetition effects from distractor faces when attending to another face under load: Evidence for a face attention resource. *Brain Research*, *1270*, 64-77.

Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104, No. 9). Englewood Cliffs, NJ: Prentice-hall.

Oberauer, K. (2019). Working memory and attention–A conceptual analysis and review. *Journal of cognition*, *2*(1).

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, *401*(6753), 584-587.

Öhman, A., Juth, P. & Lundqvist, D. (2010). Finding the face in a crowd: Relationships between distractor redundancy, target emotion, and target gender. *Cognition and Emotion*, *24*(7), 1216–1228. https://doi.org/10.1080/02699930903166882

O'Toole, A. J., Phillips, P. J., Jiang, F., Ayyad, J., Penard, N., & Abdi, H. (2007). Face recognition algorithms surpass humans matching faces over changes in illumination. *IEEE transactions on pattern analysis and machine intelligence*, *29*(9), 1642-1646.

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in cognitive sciences*, *6*(6), 261-266.

Pachella, R. G., Somers, P. Hardzinski, M. (1984). A psychophysical approach to dimensional separability. *Cognitive Psychology*, *16*(3), 279-304.

Palermo, R., & Rhodes, G. (2002). The influence of divided attention on holistic face perception. *Cognition*, *82*(3), 225-257.

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, *45*(1), 75-92.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive psychology*, *9*(4), 441-474.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. MIT press.

Pasupathy, A. (2006). Neural basis of shape representation in the primate brain. *Progress in brain research*, *154*, 293-313.

Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nature neuroscience*, *5*(12), 1332-1338.

Patching, G. R., & Quinlan, P. T. (2002). Garner and congruence effects in the speeded classification of bimodal signals. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(4), 755.

Pegna, A. J., Khateb, A., Michel, C. M., & Landis, T. (2004). Visual recognition of faces, objects, and words using degraded stimuli: where and when it occurs. *Human brain mapping*, *22*(4), 300-311.

Peelen, M. V., & Downing, P. E. (2017). Category selectivity in human visual cortex: Beyond visual object recognition. *Neuropsychologia*, *105*, 177-183.

Peelen, M. V., & Downing, P. E. (2023). Testing cognitive theories with multivariate pattern analysis of neuroimaging data. *Nature human behaviour*, *7*(9), 1430-1441.

Perry, J. R., Lupker, S. J., & Davis, C. J. (2008). An evaluation of the interactive-activation model using masked partial-word priming. *Language and Cognitive Processes*, *23*(1), 36-68.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. University of Chigaco Press

Piaget, J., & Inhelder, B. (1948). *Representation of space by the child*. Presses Universitaires de France.

Piepers, D. W. & Robbins, R. A. (2012). A Review and Clarification of the Terms “holistic,” “configural,” and “relational” in the Face Perception Literature. *Frontiers in Psychology*, *3*, 559. https://doi.org/10.3389/fpsyg.2012.00559

Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., & Kastner, S. (2009). Neural representations of faces and body parts in macaque and human cortex: a comparative FMRI study. *Journal of neurophysiology*, *101*(5), 2581-2600.

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a third visual pathway specialized for social perception. *Trends in Cognitive Sciences*, *25*(2), 100-110.

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology*, *17*(18), 1568-1573.

Pittenger, J. B., & Shaw, R. E. (1975b). Perception of relative and absolute age in facial photographs. Perception & Psychophysics, 18, 137–143.

Pomerantz, J. R. (1986). Visual form perception: An overview. *Pattern recognition by humans and machines: Visual perception*, *2*, 1-30.

Posner, M. I. (1978). *Chronometric explorations of mind*. Lawrence Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3-25.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160.

Postman, L., Bruner, J. S., & McGinnies, E. (1948). Personal values as selective factors in perception. *The journal of abnormal and social psychology*, *43*(2), 142.

Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does attention affect visual feature integration?. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(3), 361.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of neuroscience*, *18*(6), 2188-2199.

Qarooni, R., Prunty, J., Bindemann, M. & Jenkins, R. (2022). Capacity limits in face detection. *Cognition*, *228*, 105227. https://doi.org/10.1016/j.cognition.2022.105227

Quinn, K. A., & Macrae, C. N. (2005). Categorizing others: the dynamics of person construal. *Journal of Personality and Social Psychology*, *88*(3), 467.

Ramalingam, N., McManus, J. N., Li, W., & Gilbert, C. D. (2013). Top-down modulation of lateral interactions in visual cortex. *Journal of Neuroscience*, *33*(5), 1773-1789.

Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. (2016). Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior*, *58*, 354-360.

Rao, R. P. N. & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Ravizza, S. M., & Carter, C. S. (2008). Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia*, *46*(12), 2924-2935.

Reddy, L., Wilken, P., & Koch, C. (2004). Face-gender discrimination is possible in the near-absence of attention. *Journal of vision*, *4*(2), 4-4.

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.

Rees, M. (1980). The mathematical sciences and World War II. *The American Mathematical Monthly*, *87*(8), 607-621.

Richler, J. J. & Gauthier, I. (2014). A Meta-Analysis and Review of Holistic Face Processing. *Psychological Bulletin*, *140*(5), 1281–1302. https://doi.org/10.1037/a0037004

Riedl, A., & Vervaeke, J. (2022). *Rationality and Relevance Realization* (No. vymwu). Center for Open Science.

Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related fMRI investigation of implicit semantic priming. *Journal of cognitive neuroscience*, *15*(8), 1160-1175.

Rhodes, M. G. (2009). Age estimation of faces: a review. *Applied Cognitive Psychology*, *23*(1), 1–12. https://doi.org/10.1002/acp.1442

Rock, I. (1973). *Orientation and form*. New York, NY: Academic Press.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of experimental psychology: General*, *124*(2), 207.

Rossion, B., & Lochy, A. (2022). Is human face recognition lateralized to the right hemisphere due to neural competition with left-lateralized visual word recognition? A critical review. *Brain Structure and Function*, *227*(2), 599-629.

Rossion, B., Schiltz, C., Robaye, L., Pirenne, D., & Crommelinck, M. (2001). How does the brain discriminate familiar and unfamiliar faces?: a PET study of face categorical perception. *Journal of cognitive neuroscience*, *13*(7), 1019-1034.

Rushworth, M. F., Passingham, R. E., & Nobre, A. C. (2005). Components of attentional set-switching. *Experimental psychology*, *52*(2), 83-98.

Rust, N. C. & Stocker, A. A. (2010). Ambiguity and invariance: two fundamental challenges for visual processing. *Current Opinion in Neurobiology*, *20*(3), 382–388. https://doi.org/10.1016/j.conb.2010.04.013

Sanders, A. F., & Lamers, J. M. (2002). The Eriksen flanker effect revisited. *Acta Psychologica*, *109*(1), 41-56.

Sato, S., & Kawahara, J. I. (2015). Attentional capture by completely task-irrelevant faces. *Psychological research*, *79*, 523-533.

Scholnick, E.K. (1983). New Trends in Conceptual Representation: Challenges To Piaget's Theory. Psychology Press.

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. https://doi.org/10.3758/bf03207617

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), 1748–1765. https://doi.org/10.1037/0096-1523.24.6.1748

Scott, L. S., & Nelson, C. A. (2006). Featural and configural face processing in adults and infants: A behavioral and electrophysiological investigation. *Perception*, *35*(8), 1107-1128

Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., ... & Schomer, D. L. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport*, *8*(12), 2749-2754.

Semmelmann, K., & Weigelt, S. (2017). Online psychophysics: Reaction time effects in cognitive experiments. *Behavior Research Methods*, *49*, 1241-1260.

Serences, J. T., & Kastner, S. (2014). A multi-level account of selective attention. *The Oxford handbook of attention*, 76-104.

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the national academy of sciences*, *104*(15), 6424-6429.

Shalev, L., & Algom, D. (2000). Stroop and Garner effects in and out of Posner's beam: reconciling two conceptions of selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(3), 997.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379-423.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of mathematical psychology*, *1*(1), 54-87.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American philosophical society*, *106*(6), 467-482.

Soderberg, C. K., & Sherman, J. W. (2013). No face is an island: How implicit bias operates in social scenes. *Journal of Experimental Social Psychology*, *49*(2), 307-313.

Solso, R.L. (1991). *Cognitive psychology: third edition.* Allyn and Bacon

Soto, F. A., & Wasserman, E. A. (2011). Asymmetrical interactions in the perception of face identity and emotional expression are not unique to the primate visual system. *Journal of Vision*, *11*(3), 24-24.

Spruyt, A., Houwer, J. D., Hermans, D. & Eelen, P. (2007). Affective Priming of Nonaffective Semantic Categorization Responses. *Experimental Psychology*, *54*(1), 44–53. https://doi.org/10.1027/1618-3169.54.1.44

Stevens, K. A. (2012). The vision of david marr. *Perception*, *41*(9), 1061-1072.

Stolier, R. M., & Freeman, J. B. (2016). Neural pattern similarity reveals the inherent intersection of social categories. *Nature neuroscience*, *19*(6), 795-797.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643.

Stone, A., & Valentine, T. (2004). Better the devil you know? Nonconscious processing of identity and affect of famous faces. *Psychonomic Bulletin & Review*, *11*, 469-474.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in cognitive sciences*, *13*(9), 403-409.

Susac, A., Ilmoniemi, R. J., Pihko, E., Nurminen, J., & Supek, S. (2009). Early dissociation of face and object processing: A magnetoencephalographic study. *Human brain mapping*, *30*(3), 917-927.

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.

Szweigman, S. (2020). *Perceptual Specialization in Face Recognition in ASD: the Interaction Between Facial Identity and Expression* (Doctoral dissertation, University of Haifa (Israel)).

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

Teufel, C. & Nanay, B. (2017). How to (and how not to) think about top-down influences on visual perception. *Consciousness and Cognition*, *47*, 17–25. https://doi.org/10.1016/j.concog.2016.05.008

Tharmaratnam, V., Patel, M., Lowe, M. X., & Cant, J. S. (2021). Shared cognitive mechanisms involved in the processing of scene texture and scene shape. *Journal of Vision*, *21*(7), 11-11.

Theeuwes, J. (1993). Endogenous and Exogenous Control of Visual Selection. *Perception*, *23*(4), 429–440. https://doi.org/10.1068/p230429

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition*, *1*(1), 29. https://doi.org/10.5334/joc.13

Theeuwes, J., Kramer, A. F., & Atchley, P. (1999). Attentional effects on preattentive vision: spatial precues affect the detection of simple features. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(2), 341.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT press.

Thoma, V., & Lavie, N. (2013). Perceptual load effects on processing distractor faces indicate face-specific capacity limits. *Visual Cognition*, *21*(8), 1053-1076.

Thoma, V., Ward, N., & De Fockert, J. W. (2016). Misaligned and polarity-reversed faces determine face-specific capacity limits. *Frontiers in Psychology*, *7*, 1470.

Tipper, S.P. (1984). “Negative priming and visual selective attention”. University of Oxford, D.Phi1.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology Section A*, *37*(4), 571-590.

Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton University Press.

Townsend, J.T., & Ashby, F.G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), Cognitive theory. Vol. 3. (pp. 200-239). Hillsdale, N.J.: Erlbaum.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. CUP Archive.

Tranel, D., & Damasio, A. R. (1985). Knowledge without awareness: An autonomic index of facial recognition by prosopagnosics. *Science*, *228*(4706), 1453-1454.

Tranel, D., Damasio, A. R. & Damasio, H. (1988). Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology*, *38*(5), 690–690. https://doi.org/10.1212/wnl.38.5.690

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242-248.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological review*, *76*(3), 282.

Treisman, A. (1983, June). The role of attention in object perception. In *Physical and Biological Processing of Images: Proceedings of an International Symposium Organised by the Rank Prize Funds, London, England, 27–29 September, 1982* (pp. 316-325). Berlin, Heidelberg: Springer Berlin Heidelberg.

Treisman, A. (1999). Solutions to the binding problem: progress through controversy and convergence. *Neuron*, *24*(1), 105-125.

Treisman, A. M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Treisman, A., Vieira, A., & Hayes, A. (1992). Automaticity and preattentive processing. *The American journal of psychology*, 341-362.

Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences*, *24*(5), 295–300. https://doi.org/10.1016/s0166-2236(00)01814-2

Tsao, D. Y. & Livingstone, M. S. (2008). Mechanisms of Face Perception. *Annual Review of Neuroscience*, *31*(1), 411–437. https://doi.org/10.1146/annurev.neuro.30.051606.094238

Tulving, E. (1993). What is episodic memory?. *Current directions in psychological science*, *2*(3), 67-70.

Uittenhove, K., Jeanneret, S., & Vergauwe, E. (2023). From lab-testing to web-testing in cognitive research: who you test is more important than how you test. Journal of Cognition, 6(1): 13, pp. 1–17. DOI:https://doi.org/10.5334/joc.259

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & Von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological bulletin*, *138*(6), 1172.

Wallis, G. (2013). Toward a unified model of face and object recognition in the human visual system. *Frontiers in Psychology*, *4*, 497. https://doi.org/10.3389/fpsyg.2013.00497

Wang, Y., Fu, X., Johnston, R. A., & Yan, Z. (2013). Discriminability effect on Garner interference: evidence from recognition of facial identity and expression. *Frontiers in Psychology*, *4*, 943.

Wang, H., Ip, C., Fu, S., & Sun, P. (2017). Different underlying mechanisms for face emotion and gender processing during feature-selective attention: Evidence from event-related potential studies. *Neuropsychologia*, *99*, 306-313.

Wardle, S. G., Paranjape, S., Taubert, J. & Baker, C. I. (2022). Illusory faces are more likely to be perceived as male than female. *Proceedings of the National Academy of Sciences*, *119*(5), e2117413119. https://doi.org/10.1073/pnas.2117413119

Wasylyshyn, C., Verhaeghen, P., & Sliwinski, M. J. (2011). Aging and task switching: a meta-analysis. *Psychology and aging*, *26*(1), 15.

Watson, R., Latinus, M., Charest, I., Crabbe, F., & Belin, P. (2014). People-selectivity, audiovisual integration and heteromodality in the superior temporal sulcus. *cortex*, *50*, 125-136.

Weisbuch, M., & Adams Jr, R. B. (2012). The functional forecast model of emotion expression processing. *Social and Personality Psychology Compass*, *6*(7), 499-514.

Wentura, D. & Rothermund, K. (2014). Priming is not Priming is not Priming. *Social Cognition*, *32*(Supplement), 47–67. https://doi.org/10.1521/soco.2014.32.supp.47

Werner, H., & Wapner, S. (1949). Sensory-tonic field theory of perception. *Journal of personality*.

Wertz, F. J. (1983). Revolution in psychology: A case study of the new look school of perception. *Duquesne studies in phenomenological psychology*, *4*, 222-243.

Wiener, N. (1950). Cybernetics. *Bulletin of the American Academy of Arts and Sciences*, *3*(7), 2-4.

Wiese, H., Schweinberger, S. R. & Neumann, M. F. (2008). Perceiving age and gender in unfamiliar faces: Brain potential evidence for implicit and explicit person categorization. *Psychophysiology*, *45*(6), 957–969. https://doi.org/10.1111/j.1469-8986.2008.00707.x

Wiese, H., Kloth, N., Güllmar, D., Reichenbach, J. R. & Schweinberger, S. R. (2012). Perceiving age and gender in unfamiliar faces: An fMRI study on face categorization. *Brain and Cognition*, *78*(2), 163–168. https://doi.org/10.1016/j.bandc.2011.10.012

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of personality and social psychology*, *72*(2), 262.

Wheeler, J. A. (1999). Information, physics, quantum: the search for links. In *Feynman and computation: exploring the limits of computers* (pp. 309-336).

White, C. N., Ratcliff, R., & Starns, J. J. (2011). Diffusion models of the flanker task: Discrete versus gradual attentional selection. *Cognitive psychology*, *63*(4), 210-238.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic bulletin & review*, *1*, 202-238.

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: on the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 483.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it?. *Nature reviews neuroscience*, *5*(6), 495-501.

Wu, S., Hitchman, G., Tan, J., Zhao, Y., Tang, D., Wang, L., & Chen, A. (2015). The neural dynamic mechanisms of asymmetric switch costs in a combined Stroop-task-switching paradigm. *Scientific reports*, *5*(1), 10240.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of “switch costs”. *Psychological research*, *63*, 212-233.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British journal of psychology*, *79*(4), 471-491.

VanRullen, R. (2006). On second glance: Still no high-level pop-out effect for faces. *Vision research*, *46*(18), 3017-3027.

Ventura, P., Ngan, V., Pereira, A., Cruz, F., Guerreiro, J. C., Rosário, V., ... & Wong, A. (2022). The relation between holistic processing as measured by three composite tasks and face processing: A latent variable modeling approach. *Attention, Perception, & Psychophysics*, *84*(7), 2319-2334.

Victeur, Q., Huguet, P., & Silvert, L. (2019). Attentional allocation to task-irrelevant fearful faces is not automatic: Experimental evidence for the conditional hypothesis of emotional selection. *Cognition and Emotion*.

Von Bertalanffy, L. (1968). General systems theory. Penguin

Von Bertalanffy, L. (1967). General theory of systems: Application to psychology. *Social Science Information*, *6*(6), 125-136.

Xie, S. Y., Flake, J. K., & Hehman, E. (2019). Perceiver and target characteristics contribute to impression formation differently across race and gender. *Journal of personality and social psychology*, *117*(2), 364.

Yeh, Y. Y., & Eriksen, C. W. (1984). Name codes and features in the discrimination of letter forms. *Perception & Psychophysics*, *36*(3), 225-233.

Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: the role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human perception and performance*, *29*(2), 455.

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

Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. *Current Biology*, *31*(17), R1026-R1032.

Young, A. W. & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology (London, England : 1953)*, *102*(4), 959–974. https://doi.org/10.1111/j.2044-8295.2011.02045.

Young, A. W., Ellis, A. W., Flude, B. M., McWeeny, K. H., & Hay, D. C. (1986). Face–name interference. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(4), 466.

Young, A.W., Hellawell, D., and Hay, D.C. (1987). Configurational information in face perception. *Perception* 16.6 (1987): 747-7

Yuille, A. L. (1991). Deformable Templates for Face Recognition. *Journal of Cognitive Neuroscience*, *3*(1), 59–70. https://doi.org/10.1162/jocn.1991.3.1.59

Zakay, D., Bibi, A., & Algom, D. (2014). Garner interference and temporal information processing. *Acta psychologica*, *147*, 143-146.

Zebrowitz, L. A., Kikuchi, M., & Fellous, J.-M. (2010). Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. Journal of Personality and Social Psychology, 98(2), 175–189. https://doi.org/10.1037/a0017990

Zhao, M. & Hayward, W. G. (2013). Integrative processing of invariant aspects of faces: Effect of gender and race processing on identity analysis. *Journal of Vision*, *13*(1), 15–15. https://doi.org/10.1167/13.1.15

Zhou, Y., Lin, J., & Zhou, G. (2022). Misaligned dynamic faces are processed holistically. *Vision Research*, *191*, 107970.

van Zoest, W., & Donk, M. (2004). Bottom-up and top-down control in visual search. *Perception*, *33*(8), 927-937.

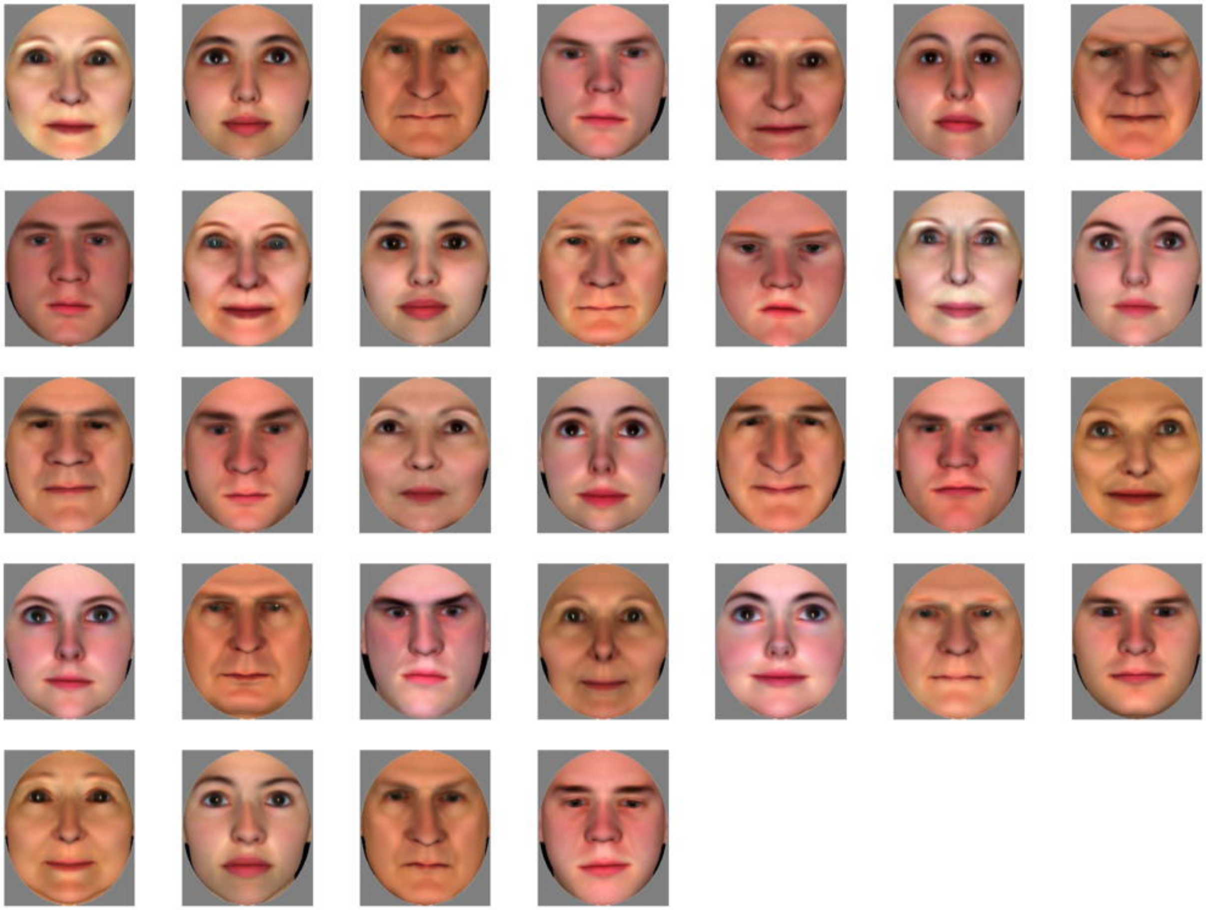
Appendices

EXPERIMENT	N	MEAN ACC
GARNER	302	0.952
ERIKSEN	360	0.954
FACE PRIME	186	0.957
SEMANTIC PRIME	176	0.962
TASK SWITCHING	184	0.956

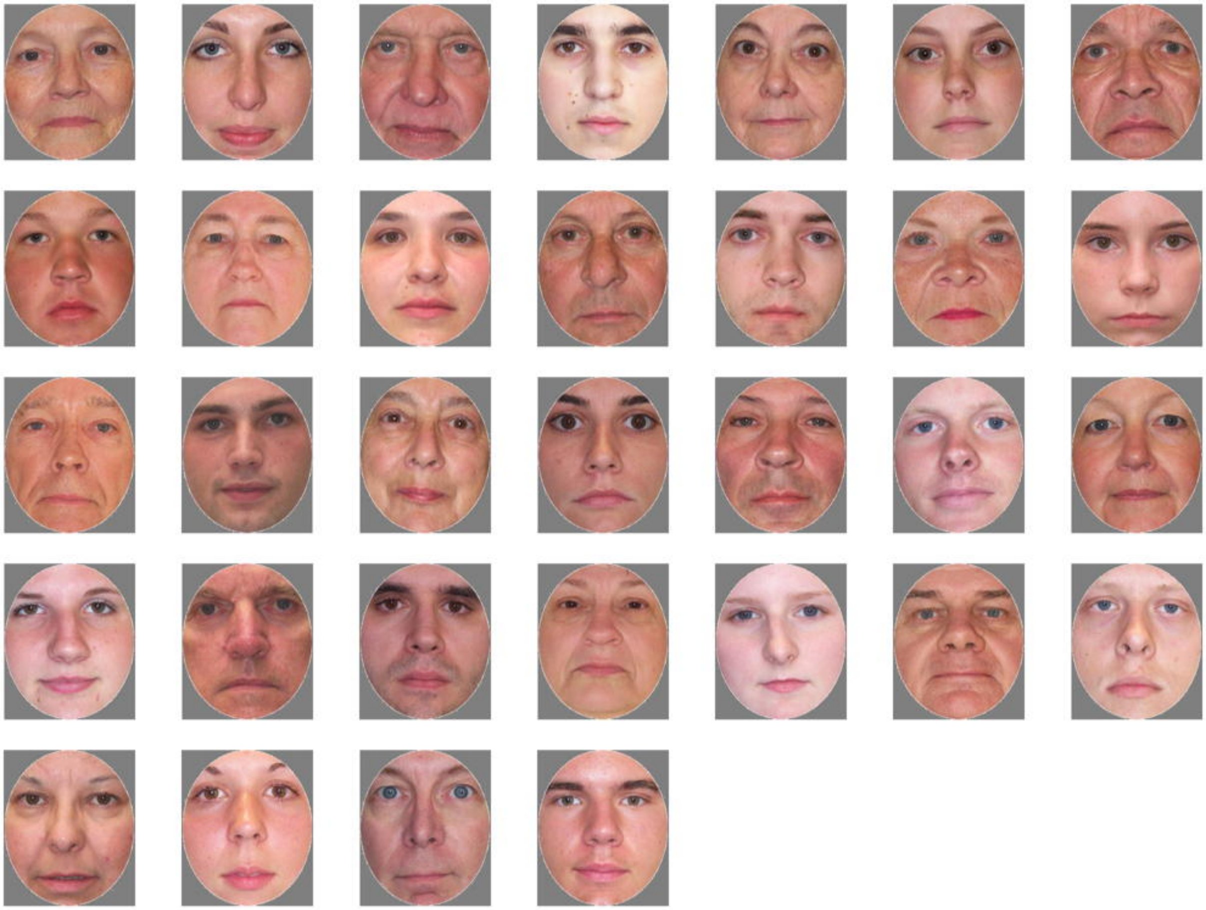
Appendix 1.1: Mean accuracy for each experiment in terms of proportion of correct responses

SET	N	MEAN ACC
FACEGEN	400	0.96
MPI	411	0.956
LIFESPAN	397	0.95

Appendix 1.2: Mean accuracy for each set in terms of proportion of correct responses



Appendix 2.1: Experimental stimuli used from Facegen *set*



Appendix 2.2: Experimental stimuli used from Lifespan *set*