Implicit Measures of Sexual Attraction

A thesis submitted for the degree of Ph.D in

Cognitive Psychology/ Neuropsychology

Leia S. H. Brasnell

School of Psychology

University of Kent

December 2023

Number of Pages: 188

Abstract

A range of implicit cognitive measures have successfully measured the sexual interest that observers hold in other people. These studies have typically focused on sexual interest in broad categories, such as images of men and women. In contrast, limited research has sought to investigate whether such measures are sensitive to the more fine-grained differences in attraction present *within* these groups. In this thesis, this was explored across a series of experiments. Firstly, sexual interest was investigated through an explicit self-report task (Chapter 2) to determine which attractiveness judgements invoke sexual interest, and thus to inform later experimental design. Then, the potential utility of implicit cognitive measurement for the investigation of sexual interest both *between* and *within* sex groups was explored (Chapter 3). The results indicate that such measures may be reflective of mate preference, whereby responses are indicative of sexual interest *between* sex groups but not *within*. The potential manipulability of implicit measures of sexual interest were also explored through the investigation of pupillary responses, by instructing observers to control their viewing behaviours and through the social presence of other people (Chapter 4). These experiments demonstrate that implicit cognitive measures of sexual interest may be susceptible to intentional control, with alternative suggestions contemplated. The findings obtained throughout this thesis are discussed in relation to previous research across a range of implicit measures and the theoretical implications considered.

Acknowledgements

I would first like to thank my supervisor Professor Markus Bindemann. Without your meticulous attention to detail, no nonsense feedback, steadfast support and understanding, I would not be half the researcher I am today.

I would also like to thank my parents for their unconditional love, support and encouragement across every aspect of my life. To my grandparents, who have believed in me every step of the way. To all of my friends across the years for the joy, laughter and companionship we have shared. To Becca with whom I shared the trials and tribulations of a BSc and a MSc. To Chloe for all the advice, care, and unwavering support. To Rachel for the emotional support and to Aidan, Jen, Bill, Nat, Lilley, Alisha and Maia, for the friendships I could not do without. Thank you to my friends from SeNSS for their encouragement and guidance, and for all the doors they have opened along the way. I would also like to thank the entire tech team for saving me and my thesis too many times to count.

Special thanks to my fiancée Emily for being there every step of the way. Thank you for the late nights listening to me ramble on about my work. For the endless support, love and motivation at both the high and low points of this PhD. For navigating a world-wide pandemic, grief and growth by my side. I could not have done this without you.

Last but not least, the most serious thank you of them all. Thank you to Gary, Switchboard and Chiron, my cats. Thank you for all the comfort, joy, love and, at times, chaos you have brought to my world.

Declaration

I declare that this thesis is my own work carried out under the normal terms of supervision.

____

Leia S. H. Brasnell

Table of Contents

Chapter 1

General Introduction

1.1 Background

Humans form attractiveness evaluations of others rapidly (Locher et al., 1993; South Palomares & Young, 2018), even when people are only viewed briefly (Olson & Marshuetz, 2005). These appraisals occur from infancy (Hoss & Langlois, 2003; Langlois et al., 1991; Slater et al., 1998) and have clear societal impact with more attractive candidates preferred when hiring for jobs or in voting situations (Efrain & Patterson, 1974; Frieze et al., 1991; Marlowe et al., 1996). As such, attraction has become a prolific avenue of research, particularly through investigation of observers' sexual interest in other people.

Within Psychology, sexual interest has been investigated with a range of measures, beginning with self-report (Dombert et al., 2016; Hoon et al., 1976; Howell et al., 1987). However, this relies on genuine response and honest reactions, which poses an inherent flaw when considering the sensitive nature of sexual interest. Therefore, research in this field has re-oriented towards identifying alternate measures that provide objective results, such as genital responses (Barbaree et al., 1979; Rellini et al., 2005; Suschinsky et al., 2009). However, these methods are highly invasive and may also be at risk of manipulation (Mahoney & Strassberg, 1991), providing a need for alternative approaches.

Consequently, cognitive tasks have been developed to combat these issues. Measures such as viewing time (Ebsworth & Lalumière, 2012; Israel & Strassberg, 2009; Mokros et al., 2013), reaction time (Santtila et al., 2009; Wright & Adams, 1999), Stroop tasks (Ciardha & Gormley, 2012; Mannfolk et al., 2023; Thornton & Laws, 2009) and pupillary responses (Attard-Johnson & Bindemann, 2017; Rieger & Savin-Williams, 2012; Watts et al., 2017), have all been used for investigating sexual interest. Generally, such attraction is evaluated by comparing gender categories, with very limited exploration within these groups (Israel & Strassberg, 2009; Lippa, 2012; Lippa et al., 2010). As such, the sensitivity of these responses in evaluating attractiveness at the fine-grained level within gender category is largely unknown and investigated. This thesis aims to explore the utility of these cognitive measures for analysing sexual interest within gender categories.

This chapter begins with an evaluation of the measures that are currently used to investigate sexual attraction. I will first discuss genital response measures, and then consider the increasingly employed cognitive measures of viewing time, Stroop tasks and pupillary responses. I will then discuss the strengths and weaknesses of such tasks. Next, I will evaluate existing research surrounding the ability of cognitive measures in assessing attractiveness responses within gender categories. Finally, I will discuss some of the factors which may interact with sexual interest, focusing specifically on familiarity. I will end this chapter by outlining the empirical research contained within this thesis.

1.2 Implicit Measures of Attraction

Historically, sexual interest has been explored using self-report (Dombert et al., 2016; Hoon et al., 1976; Howell et al., 1987). For example, participants are often asked to indicate their sexual orientation according to the Kinsey scale (Kinsey et al., 1948), a 7-point scale designed based on sexual experience and desire. Possible responses range from 0 indicating *'exclusively heterosexual'* to 6 for *'exclusively homosexual'*. However, this measure contains an inherent flaw by requiring honest response, undermining the ability to collect accurate data when considering a concept as culturally sensitive as sexual interest (Meston et al., 1998). In an attempt to address this caveat, more objective cognitive methods have been developed.

1.2.1 GENITAL RESPONSE

In an attempt to combat the problems of limited validity, research investigating sexual interest has moved to analysing unconscious physiological responses to sexual stimuli. To this effect, the plethysmograph has been employed to analyse genital arousal response (Barbaree et al., 1979; Barker & Howell, 1992; Müller et al., 2014). This measure involves placing a strain gauge around the penis in order to calculate circumference as an indicator of arousal. An alternative measure of photoplethysmography is also used in which a device is inserted into the vagina and arousal evaluated by blood volume within the vaginal walls. Much of the research in the field of sexual interest has adopted this measure, with claims that it represents the optimal standard for sexual interest research (Murphy et al., 2015). However, this method also has its limitations. The plethysmograph involves highly intrusive technology that requires placement around the genitalia of men, or insertion into the genitalia of women, to record data. This severely limits the accessibility of this research as the participant pool is reduced due to the invasive nature of this method. Consequently, the remaining participants are likely to be more sexually experienced and confident (Wolchik et al., 1985), thus reducing ecological validity. Furthermore, evidence suggests that even this coveted measure contains an element of fakeability, whereby participants intentionally suppress or enhance their genital arousal, indicating that this response may not assess sexual interest as directly as previously believed (Kalmus & Beech, 2005; Trottier et al., 2014). As such, research has shifted away from dependency on this response, and towards determining alternative accessible measures able to directly evaluate sexual interest.

1.2.2 VIEWING TIME

The development of cognitive latency-based measures has led to an increase in popularity of non-invasive implicit measures of sexual interest, such as viewing time. Originally employed by Rosenzweig (1942), viewing time tasks operate under the assumption that attentional delays are caused by sexually interesting images (Gress et al., 2013). These stimuli act as a distractor within cognitive tasks as they engage observers' attention (Wright & Adams, 1999). As such, stimuli featuring models congruent with participant sexual orientation will be viewed for longer than those with a non-interesting other (see Figure 1.1). For example, heterosexual males will look longer at female than male stimuli.

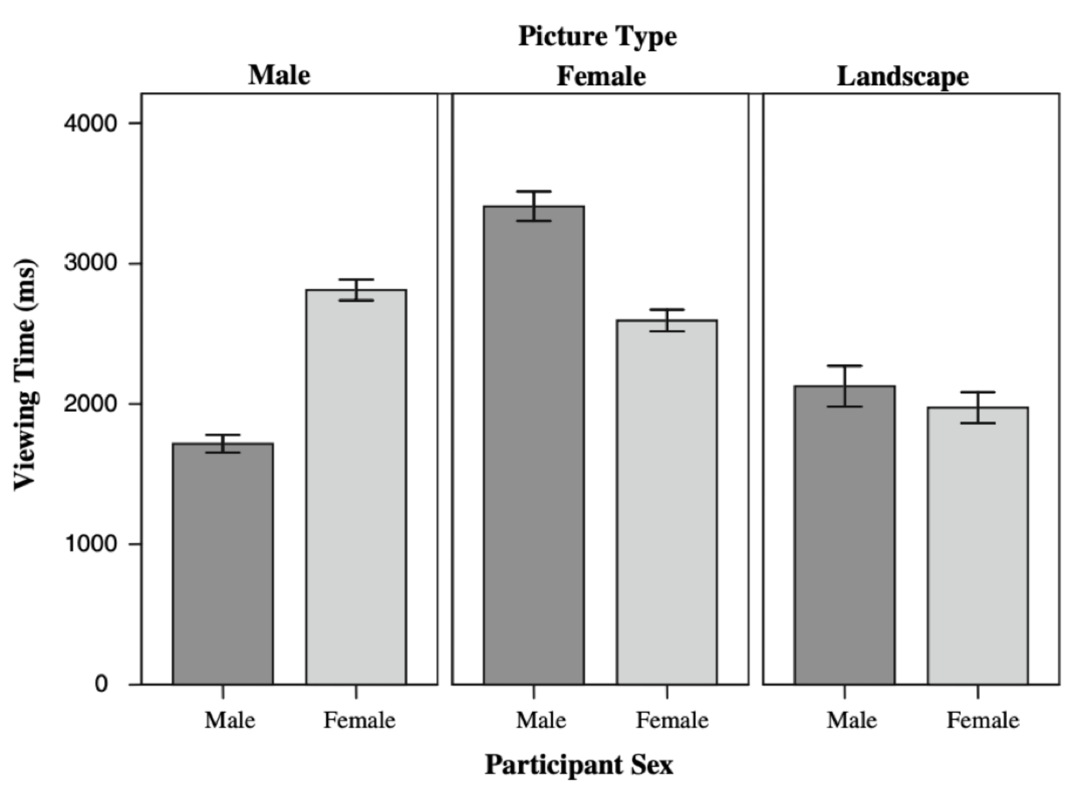

Figure 1.1. An example of viewing time data gathered when heterosexual participants view male, female and neutral stimuli (Taken from Israel & Strassberg, 2009).

This effect has been found across a multitude of experiments (Ebsworth & Lalumière, 2012; Imhoff et al., 2010; Mokros et al., 2013), with the consistency of results highlighting the robust nature of this measure. Furthermore, this measure has been successfully employed within forensic populations, (Gray et al., 2015; Gress, 2005; Harris et al., 1996), whereby deviant attraction, i.e. sexual interest in children and/or violent sex offenses, has been accurately identified. As such, viewing time appears to be an unintrusive and effective measure of sexual interest.

1.2.3 STROOP TASK

The Stroop task was developed in the early twentieth century (Stroop, 1935), and has since seen extensive experimental use (Epp et al., 2012; MacLeod, 1991). In this task, colour names are presented in various hues and participants are instructed to indicate word tint rather than the written name. When name and hue are incongruent, an interference effect causes longer reaction times. This finding is corroborated by anterior cingulate cortex activation during incongruent trials (Hanslmayr et al., 2008; Hayward et al., 2004; Pardo et al., 1990), which is an area of the brain often thought to be utilised during task interference (Haupt et al., 2009; Taylor et al., 1994; van Veen et al., 2001).

More recently, adaptations of the traditional Stroop have been developed, for example the Emotional Stroop replaces colour names with emotional and neutral words, with tint indication time slowed for emotional words (Ashley & Swick, 2009; McKenna & Sharma, 1995; Watts et al., 1986). This emotional interference is particularly prevalent within clinical populations when words reflect themes that are pathologically salient for participants (Williams et al., 1996). This effect has been uncovered within a multitude of populations, for example, individuals with a diagnosis of depression (Epp et al., 2012; Mitterschiffthaler et al., 2008; Williams & Nulty, 1986), posttraumatic stress disorder (Cisler et al., 2011; Khanna et al., 2017; Metzger et al., 1997), panic disorder (Dresler, Attar, et al., 2012; Dresler, Ehlis, et al., 2012) and anxiety (Becker et al., 2001; Dresler et al., 2009; Richards et al., 1992). The consistency of these findings indicates the robust nature of the Emotional Stroop task.

With the established validity of the Emotional Stroop, this task has been extended to investigate sexual interest. Here, words high in sexual content are presented and reaction times mirror the strength of attraction through attentional bias (Thornton & Laws, 2009). Much of the research in this area has been undertaken within forensic populations, such as sexual offenders, (Price et al., 2012, 2013; Spada & Jeglic, 2016) to investigate deviant sexual interests such as sexual abuse or pedophilia. While this measure has seen success in differentiating between offending and non-offending populations, conflicting evidence remains surrounding whether differences within deviant interests can be identified (Price et al., 2013; Smith & Waterman, 2004).

The Emotional Stroop task has been further adapted into the Pictorial Stroop task, whereby images are presented with a colour tint rather than words. This provides a more ecologically valid approach for investigating attentional interference, as the stimuli are representative of real-life experience and provide more salient information (Constantine et al., 2001; Lavy & van den Hout, 1993). This is particularly poignant when considering paedophilic sexual interest, which is often analysed using the Tanner stages (Ebsworth & Lalumière, 2012; Mokros et al., 2010). The Tanner stages refer to the five stages of sexual development in accordance to secondary sex characteristics, such as pubic hair and genital size (Tanner, 1973). The differences in these stages are more easily distinguished using visual stimuli than in lexical format. Furthermore, even when considering non-age-related interest, images allow for the inclusion of more sexually salient information than can be inferred from words. Indeed, the Pictorial Stroop has been utilised to effectively indicate sexual interest across a range of experiments, with longer naming times for stimuli comprised of people congruent with the observer's sexual orientation (Ciardha & Gormley, 2012; Mannfolk et al., 2023; Ó Ciardha & Gormley, 2013).

1.2.4 PUPILLARY RESPONSE

Whilst measures such as viewing time and Stroop tasks can measure implicit cognition, these tasks remain open to the possibility of conscious manipulation (Akerman & Beech, 2012). For example, if an individual was motivated to disguise their true responses, then they could alter their behaviour in these paradigms by intentionally altering their response patterns, consequently skewing their data. The measurement of pupil diameter has been touted as an effective solution to this issue.

The pupils of the eye change in size to regulate the amount of light that hits the retina. Pupil size is controlled by two muscles, influenced by the parasympathetic and the sympathetic branches of the autonomic nervous system. The sphincter pupillae is primarily governed by the parasympathetic nervous system and is responsible for constriction, whilst the dilator pupillae is largely controlled by the sympathetic nervous system and dictates pupillary dilation. The collective influence of these muscles determines pupil size (McDougal & Gamlin, 2015; Zele & Gamlin, 2020). As such, pupillary responses are considered autonomic and impervious to conscious manipulation.

However, recent developments in research indicate that pupil size change is a multifaceted response. Pupillary responses have been found to be reflective of a number of factors, such as luminance (Guillon et al., 2016; Kun et al., 2012; Watson & Yellott, 2012); cognitive load (Matthews et al., 1991; Vogels et al., 2018; Zekveld & Kramer, 2014), and arousal (Kim et al., 2000; Matthews et al., 2020; McGarrigle et al., 2017).

One such factor is emotional affect. Pupillary responses have been extensively employed to investigate this response. Generally, research supports the idea that increases in the size of the pupils of the eye are indicative of emotional affect (Janisse, 1973; Oliva & Anikin, 2018; Snowden et al., 2016). Such research has shown that when presented with auditory stimuli, such as couples fighting (Partala & Surakka, 2003; Rosa et al., 2017), and visual stimuli, such as images of babies (Chen et al., 2017; Dietz et al., 2011), pupil dilation increases as a reflection of intensity of emotional content. This measure has also seen some success in assessing emotional processing within populations considered to exhibit a deficit in this area, such as individuals with Parkinson's disease (Dietz et al., 2011; Schwartz et al., 2018) or psychopathy (Burley et al., 2017, 2019). For example, individuals with Parkinson's disease elicit greater pupil size when viewing emotional images, such as violence, compared to neutral images, such as office buildings (Dietz et al., 2011). This suggests that pupillary responses are particularly effective for analysing emotional affect, as they are sensitive to such affect even within populations experiencing a deficit in emotional functioning. This response is robust and can even become elevated under conditions such as sleep deprivation (Franzen et al., 2008, 2009) and depression (Siegle et al., 2003; Steidtmann et al., 2010).

The nature of pupillary responses to emotional affect appear to reflect intensity of emotion, rather than valence. In one study, which investigated pupil size changes whilst participants viewed emotional images featuring person stimuli, increasing pupil dilation was observed during presentation of facial images exhibiting both positive and negative expressions (Lang et al., 1997) when compared to neutral stimuli (see Figure 1.2) (Bradley et al., 2008). This suggests that pupil size indicates strong emotional affect, regardless of type of emotion. Furthermore, skin conductance responses covaried with pupil size, indicating that this response is sensitive to stimulation of the sympathetic nervous system.

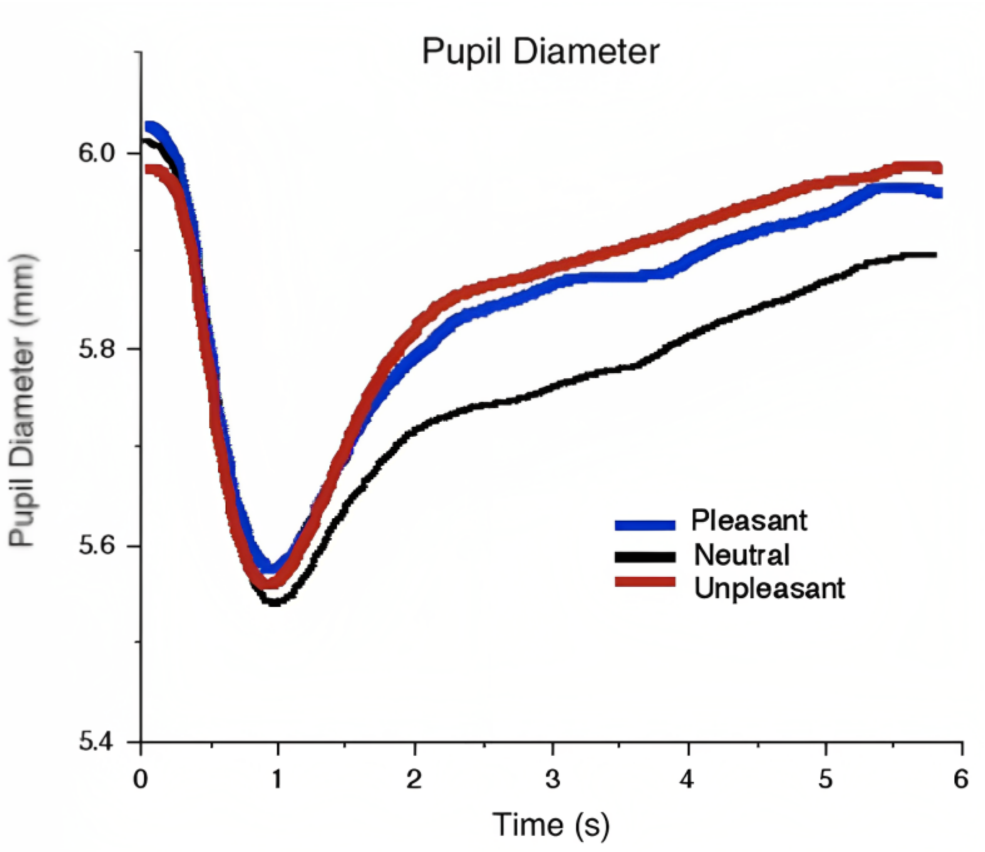

Figure 1.2. Graph showing the greater pupillary increases for emotional words following the initial light reflect (Taken from Bradley et al., 2008).

Experiments assessing pupil dilation to emotional affect have consistently provided evidence that this response is sensitive to feelings elicited within participants. As such this measure may also be suitable to effectively measure sexual interest. The stimuli utilised for the assessment of attraction often feature images of people in various states of undress. These images can be considered inherently emotional, as they are designed to elicit the positive emotions surrounding sexual arousal. Furthermore, the social content of the stimuli also suggests that pupil size is a potential measure for sexual interest. Research has suggested that oxytocin can increase pupil size in order to heighten attention to socially relevant stimuli (Leknes et al., 2013; Prehn et al., 2013). This has been demonstrated through the use of social (faces) and non-social (clocks) images initially presented outside of awareness (Xue et al., 2020). Consequently, this data suggests that oxytocin promotes allocation of attention towards social stimuli, simultaneously eliciting pupil dilation. Sexual stimuli and person images are intrinsically connected and inseparable, and as such it seems plausible that pupillary response may also be indicative of sexual interest.

Indeed, the relationship between pupil size and sexual attraction has fascinated researchers for the last six decades. Pupil size has been proposed as a potentially useful measure of sexual interest (Attard-Johnson & Bindemann, 2017; Hess et al., 1965; Rieger et al., 2015), whereby sexual orientation can be alluded to through increase in the size of the pupils of the eye. Pupillary responses were first explored in the 1960's (Hess et al., 1965; Hess & Polt, 1960), where preliminary findings discovered that the pupil reacts to sexual stimuli. Specifically, the pupil dilates more to sexual stimuli congruent with sexual orientation than stimuli not aligned with sexual preference. The analyses of pupil size in these early experiments were conducted by photographing participant's eyes whilst viewing the stimuli and measuring diameter manually with a ruler. Consequently, whilst this research showed promise for pupillary responses as a measure of sexual interest, the findings were limited by the primitive technology available at the time. As such, subsequent research utilising this method was infrequent (Atwood & Howell, 1971; Green et al., 1979;

Hamel, 1974). However, with the development of more sophisticated eye-tracking technology, this measure has increased in popularity (Attard-Johnson & Bindemann, 2017; Snowden et al., 2019; Watts et al., 2017).

Recent evidence using eye-tracking technology has confirmed that the pupil dilates when observers are presented with sexual stimuli reflective of their sexual orientation. For example, in an experiment conducted by Rieger and Savin-Williams (2012), pupil size was recorded whilst participants viewed videos featuring male and females masturbating. Pupil size increased during the viewing of sexual stimuli that were congruent with sexual orientation, with consistent pupil size patterns found across all stimuli in each sex category, highlighting the reliability of this relationship. Additionally, pupil dilation correlated with self-report and viewing time measures, highlighting the validity of this measure.

Whilst evidence generally supports the idea that pupillary responses are an effective autonomic measure of sexual interest (Attard-Johnson et al., 2021; Rieger et al., 2015; Rieger & Savin-Williams, 2012), there is some evidence that top-down control can be exerted over pupil size through application of cognitive techniques, such as the imagining of past emotional experiences (Ehlers et al., 2016, 2015). However, the pupillary changes caused by these techniques were brief and showed varying degrees of success. Furthermore, these experiments investigated conscious pupil control absent of other stimuli or tasks. Therefore, the generalisability of these findings is uncertain, as it remains unknown if these techniques can interfere with pupillary responses observed in response to factors such as sexual interest.

Additionally, wavering attention has been correlated with diminished pupil size (Brink et al., 2016; Grandchamp et al., 2014). As such, if a participant were to intentionally decrease attention to a task or stimuli, it is plausible that pupil size would be reduced, which can either add noise to the data or may be misinterpreted as a reflection of the task or factor being investigated. However, this can be prevented by using a different experimental design. For example, when analysing sexual interest, evidence has shown that attention can be unconsciously captured by explicit images that are presented subliminally (Jiang et al., 2006). Therefore, even if attention is later diverted or diminished, initial pupil size should reflect true autonomic response. Consequently, despite some evidence suggesting that pupil size can be intentionally manipulated, the conditions under which this may occur are restrictive and under-researched. As such, the adoption of this measure for investigating cognitive factors such as sexual interest remains promising.

1.3 Strengths and Limitations of Cognitive Measures of Sexual Interest

Despite the clear merits of employing cognitive measures for investigating sexual interest, there are a number of factors that complicate the application of these measures.

1.3.1 GENDER-BASED RESPONSES

Male and female observers respond differently to sexual stimuli. Research has proposed that while male sexual responses have evolved to develop sexual arousal to specific targets, female sexuality is more fluid and less oriented toward particular individuals, resulting in reduced physiological reactions towards specific targets (Bailey, 2009; Rieger et al., 2015).

Such differences have been observed across a range of cognitive measures. Evidence suggests that the relationship between genital arousal and self-report is weaker in women than men (Chivers et al., 2007) and that females exhibit a generalised response to sexual content rather than to a specific gender (Chivers et al., 2010; Peterson et al., 2010). This is supported, for example, by data showing that women exhibit genital arousal even to animal scenes (Chivers & Bailey, 2005). Furthermore, viewing time measures mirror this gender difference in response (Lippa et al., 2010; Quinsey et al., 1996; Xu et al., 2017), highlighting this robust multi-measure finding.

When considering pupillary response as a measure of sexual interest, the clear patterns of pupil size correspondence to sexual orientation have only been noted in male participants (Attard-Johnson et al., 2016; Rieger et al., 2015). Whilst some studies have found congruent pupil dilation and self-reported sexual interest in female observers (Attard-Johnson & Bindemann, 2017; Hamel, 1974), many have found this link to be more obscure. For example, in many cases heterosexual and bisexual women elicit increased pupil response to images of both men and women (Attard-Johnson et al., 2016; Rieger et al., 2016; Rieger & Savin-Williams, 2012).

It remains unknown what governs these gender differences within cognitive measures of sexual interest. There has been some suggestion that differences in attention may drive these responses. Eye-tracking experiments show variations in gaze across men and women whereby males fixate on female stimuli faster and view them for longer instances of time than females. On the other hand, women fixated comparably quickly on both male and female images, and in some cases orientated towards female images faster than those of men (Dawson & Chivers, 2016). Furthermore, when actually observing these stimuli, male and female participants often show different viewing patterns. For example, male observers spend more time looking at female faces than male, with female participants viewing clothing and genital regions for longer than men (Rupp & Wallen, 2007). These differences may contribute to the presence of gender differences in cognitive measurement of sexual arousal.

The reason as to why there are consistent gender differences in sexual response remains mostly unknown. A leading theory is grounded in evolutionary psychology and referred to as the preparation hypothesis (Suschinsky & Lalumière, 2011). When any sexual cue is perceived, vaginal blood flow and lubrication occurs. It is theorised that this response has evolved to protect and reduce damage to reproductive organs in cases of non-consensual sexual activity. It has been documented that women (and non-female identifying individuals with vaginas) experience physical arousal to non-consensual stimuli (Laan et al., 1995; Suschinsky & Lalumière, 2011) and real-world experience (Levin & van Berlo, 2004). As these groups have been exposed to the threat of sexual violence across human history (Petrak, 2002; Thornhill & Palmer, 2001), it is possible that gender differences may be due to evolutionary pressures. However, further research is required to continue the exploration into potential reasons for the differences in sexual response across genders before any conclusions can be confidently drawn.

1.3.2 AGE-RELATED SEXUAL INTEREST

As discussed, cognitive measures of sexual interest can be effective when considering sexual orientation, especially within male participants. However, a second factor of importance within the study of sexual interest is a preference for different ages. Genital responses are generally considered effective in differentiating between age groups (McPhail et al., 2019), and have been employed for use within forensic populations and environments (Clift et al., 2009; Müller et al., 2014; Murphy et al., 2015). However, this measure is highly invasive and requires specialist equipment.

Viewing time procedures have been proposed as a promising alternative, with data showing this response can effectively discriminate between attraction to different age groups, i.e. adults and children (Abel et al., 1998, 2004; Schmidt et al., 2017). Pictorial Stroop tasks have also found success when distinguishing between age groups (Ciardha & Gormley, 2012; Ó Ciardha & Gormley, 2013), although these papers represent the extent of empirical evidence for this measure. Further research will be required to make firm conclusions on the applications of this response.

Pupil dilation measures offer a non-intrusive, objective measure of sexual interest, Evidence has shown reliable support for pupil size as a measure of age-related sexual interest. In an experiment by Attard-Johnson et al., (2016), pupil dilation was measured when viewing males, females and children in swimsuits (see Figure 1.3), a common design choice (Attard-Johnson et al., 2017, 2019; Ó Ciardha et al., 2018), based on findings suggesting that sexual interest can be examined using non-explicit imagery. Results from this experiment showed that both male and female participants experienced pupil dilation specific to viewing stimuli featuring adults, as opposed to children. Furthermore, consistent with previous findings, male participants dilated to female stimuli, whilst female participants exhibited increased pupil size to both male and female images. Therefore, evidence suggests that pupillary responses are sensitive to age-related sexual interest in both males and females, as well as general sexual interest in males.

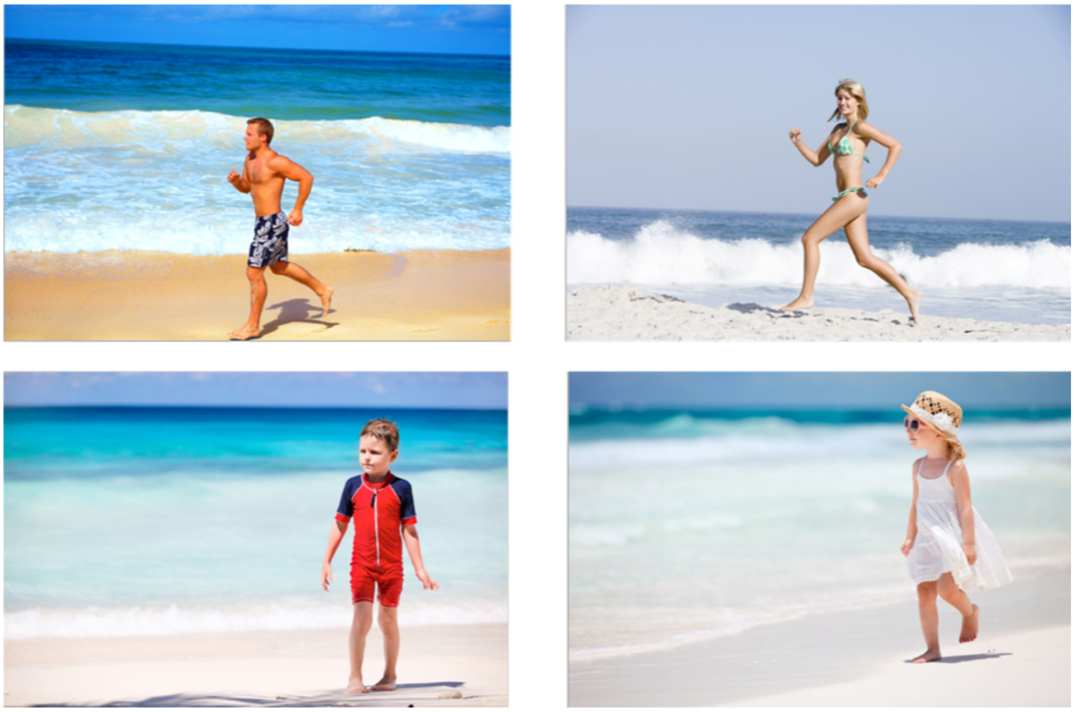

Figure 1.3. Examples of non-explicit stimuli utilised to investigate pupillary responses as a measure of age-specific sexual interest (Taken from Attard-Johnson et al., 2016).

1.3.3 SEXUAL EXPLICITNESS OF STIMULUS MATERIALS

When considering the measurement of sexual interest, a wide range of stimuli have been employed from clothed (Israel & Strassberg, 2009; Pezzoli et al., 2022; Rullo et al., 2015) and naked (Fromberger et al., 2015; Harris et al., 1996; Morandini et al., 2019) to mixed stimuli types (Dombert et al., 2013; Ebsworth & Lalumière, 2012; Gress, 2005). A trend of increased interference with attention has been identified within viewing time tasks, whereby the more sexually explicit the stimuli are, the longer they are viewed (Brown et al., 1973). However, evidence suggests that this relationship may instead be curvilinear, with viewing times decreasing when pornography becomes highly erotic (Love et al., 1976). Some evidence suggests that while sexually explicit imagery leads to the most disruptions when performing a cognitive task, clothed stimuli can also elicit longer delays than neutral imagery (Schmidt et al., 2017; Wright & Adams, 1999).

The effect of stimulus type on cognitive measures of sexual interest has also been considered within pupillary responses. Pornography has been employed in many instances when assessing sexual interest in response to pupil size (Rieger et al., 2015; Rieger & Savin-Williams, 2012), however the nature of these stimuli reduces the potential applications. Akin to other measures, explicit images and videos appear to cause generally greater pupil dilation than non-explicit stimuli, however these stimuli can also be effective in producing sexual interest-based responses (Watts et al., 2017). In this experiment, pupil size was measured when observers were presented with pornographic material or people discussing the weather. Both explicit and non-explicit stimuli produced pupil dilation congruent with sexual orientation. However, these effects were less pronounced in the non-explicit stimuli condition, suggesting that pupillary responses may be an effective measure of sexual interest when utilising non-invasive design, albeit to a lesser extent than when employing more explicit stimuli. It should be noted that in this experiment the stimuli were not matched for identity across explicitness conditions. Therefore, individual differences in participant attraction to specific identities may have been present, potentially confounding results related to stimuli explicitness.

The role of identity confounds when exploring stimuli explicitness and pupillary responses of sexual interest have been addressed in an experiment by Attard-Johnson and Bindemann (2017), which explored pupillary responses to naked, blurred and clothed images of identity matched individuals (see Figure 1.4). It was found that dilation was comparable for naked and dressed targets and corresponded with self-reported sexual orientation. Thus, suggesting the possibility that pupillary responses indicate sexual interest without being specific to explicit sexual content. However, this paper currently stands alone in investigating stimuli content whilst controlling for identity. As such, further research is necessary for solid conclusions to be drawn.

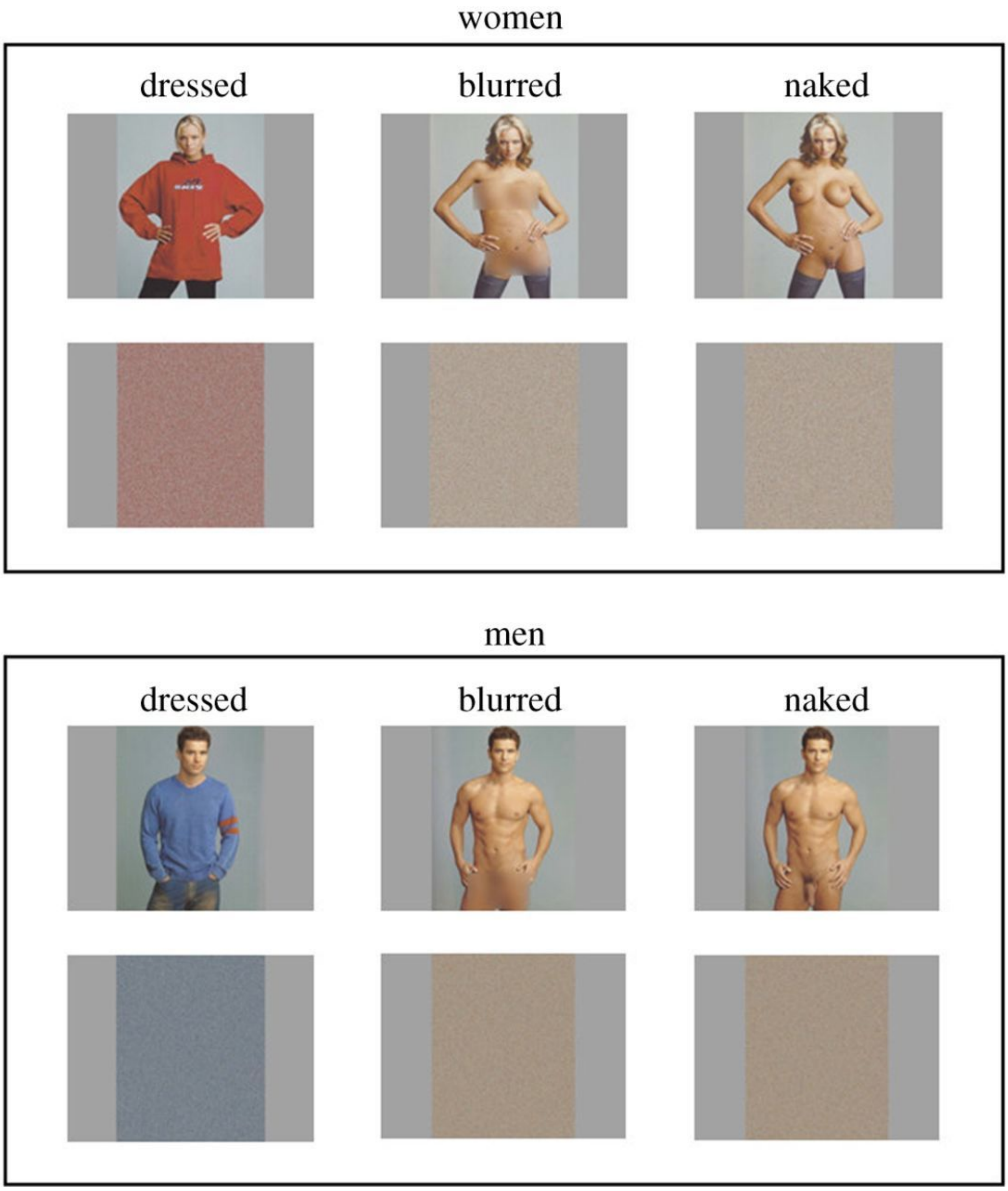

Figure 1.4. Example of the stimuli utilised in Attard-Johnson and Bindemann (2017) to investigate pupillary responses to explicit and non-explicit images.

1.4 Within-Category Attraction

Thus far, the vast majority of experiments have investigated sexual interest across gender category, representing sexual orientation. However, research into attraction *within* these groups has been limited, thus it remains uncertain whether cognitive measures of sexual interest can differentiate between attraction within-gender categories. One paper examined attraction both across and within category, by presenting participants with images of partially clothed men and women taken from popular magazines (Israel & Strassberg, 2009). Viewing times mirrored that found across the literature, with heterosexual participants viewing opposite sex images longer than same sex. Within-category effects were investigated by correlating self-reports of sexual appeal and viewing times for specific images (i.e. on a by-item basis), but no relationship was found. This experiment suggests that while implicit cognitive measures of sexual interest can be effective in analysing attraction across gender categories, the potential applications of this response within these groups might be limited.

However, there is also some indication that cognitive tasks are sensitive to within-group differences in attraction. In an experiment where participants viewed models in swimwear, viewing time increased with stimulus attractiveness (see Figure 1.5) (Lippa, 2012). For heterosexual men, this was specific for images of females, whilst women exhibited this pattern regardless of image gender. This pattern of category specificity within male participants was replicated within homosexual males, with longer viewing times for male stimuli. However, viewing time differences dependent on attraction within sex categories were less pronounced for homosexual in comparison to heterosexual males. Heterosexual women showed comparatively weak category specificity, with homosexual women tending towards sex category specificity, albeit to a lesser extent than male participants. These results converge with previous findings detailing similar responses in an almost identical paradigm in which only heterosexual participants were considered (Lippa et al., 2010). These two papers taken together suggest the possibility that cognitive measures may be sensitive to attraction within gender categories. However, due to the limited amount of research surrounding these responses and the fact that these analyses are often secondary hypotheses, it remains unclear to what extent this may be the case. Consequently, further research prioritising the evaluation of within-category cognitive measures of attraction is required.

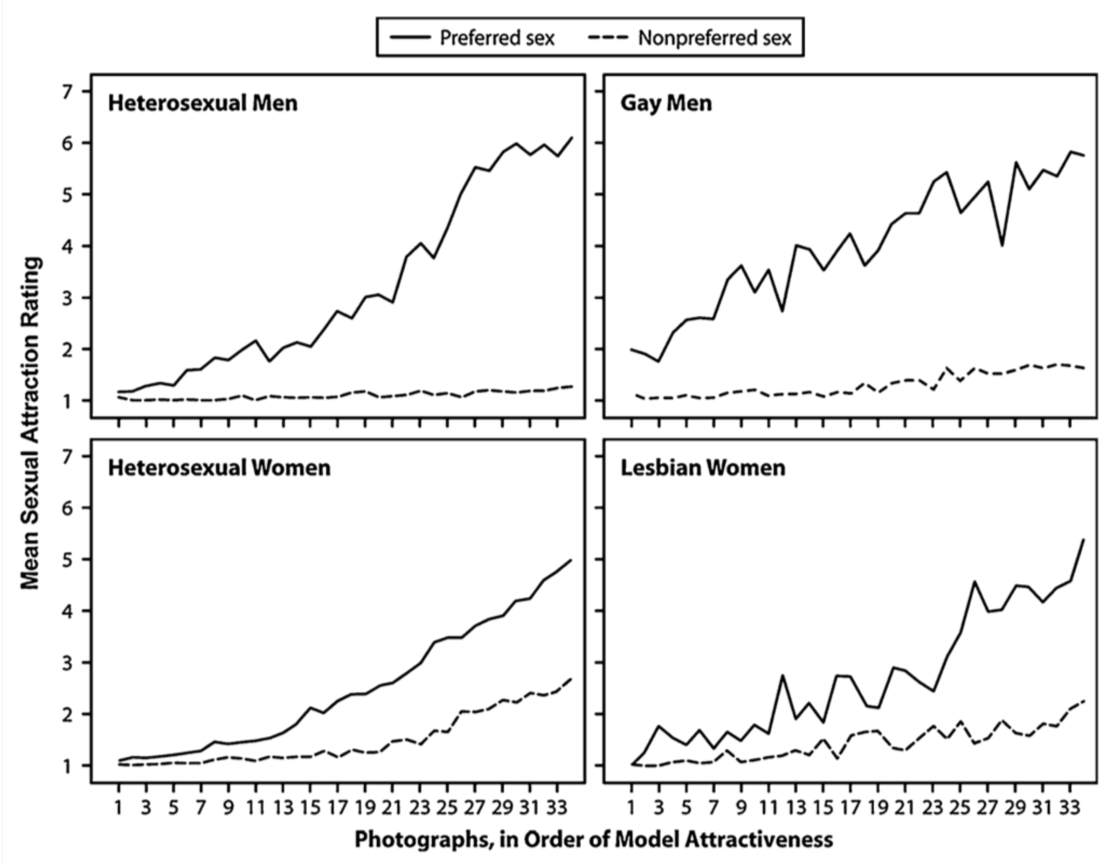

Figure 1.5. Array of graphs taken from Lippa (2012) detailing viewing times and sexual attractiveness ratings of stimuli exhibited by heterosexual and homosexual men and women.

Within these experiments, attractiveness is often examined across stimuli featuring different image models. However, this leaves open the potentiality for identity based confounds, such as individual differences in attraction. As such, a more reliable measure would be to use multiple images of a single person, differing in attractiveness. Individuals can differ substantially in appearance across images (Burton et al., 2016; Burton, 2013). This within-person variability can be so large that two different individuals may appear more similar in appearance than the same person in two different photographs (Andrews et al., 2015; Kramer et al., 2018). However, this within-person variation is also functional. For example, it can facilitate the robust learning of a person's facial identity, by capturing their appearance across a range of conditions (Baker et al., 2017; Murphy et al., 2015; Ritchie & Burton, 2017). Thus, it is clear both that appearance can vary within-person, and that this variation is key to identity. As such, it stands to reason that the same can be expected of attractiveness.

To examine this, Jenkins et al. (2011) developed a Card Sorting Task in which participants were given 20 images featuring two separate unfamiliar people. When asked to categorise these images by identity, a median of 7.5 groups were formed. As such, when observers were given a stimulus set which included images of only two identities, participants were poor at categorising identity, instead believing there were considerably more identities present than there were. Furthermore, participants provided attractiveness ratings across different images of the same face. Variation within this factor was extensive, with attractiveness ratings differing more within an identity than between separate people. The suggestion that attractiveness varies considerably across identity has also been supported through a number of other experiments where facial stimuli are examined when considering differences in facial expression and make-up (Jones & Kramer, 2015; Sutherland et al., 2017). The sum of this evidence indicates that attractiveness is not a stable attribute within identity. This suggests that multiple images of the same individual can be utilised to manipulate attractiveness while controlling for identity based confounds, thus facilitating the measurement of sexual interest through cognitive measures.

1.5 The Familiarity Caveat

1.5.1 FAMILIARITY

When utilising cognitive methodology for the assessment of sexual interest, it is important to consider any other factors that may account for or confound response. The familiarity of visual stimuli has been found to exert influence over such measures (Greene & Rayner, 2001; Lancry-Dayan et al., 2018). For example, familiarity is another factor that has been found to influence the pupil response (Kafkas & Montaldi, 2015; Naber et al., 2013; Weiss et al., 2016). The pupil old/ new effect (Võ et al., 2008) has been observed consistently in psychological research and refers to the increased pupil dilation present when viewing familiar, or previously viewed, stimuli. For example, this effect was explored by Kafkas and Montaldi (2015). They exposed participants to 200 images in grayscale in a randomised two task design. Half of these stimuli had been previously presented to participants, and thus encoded into memory. Greater pupillary responses were detected for familiar than novel stimuli.

Interestingly, evidence has found that this response may be modulated by emotional content. Võ et al. (2008) discovered that this old/ new effect was greatest for words considered neutral, and significantly reduced for words with emotional content (see Figure 1.6). In this experiment, the old/ new effect was categorised by greater pupillary dilation to hits than correct rejections, which are less cognitively demanding. The authors propose that the effort required to correctly categorise an emotional word as new is higher than that of neutral words, contributing to the diminished effect for emotional words. As such, it appears that factors that influence pupil dilation, such as familiarity and emotional affect, may interact to determine pupillary responses. Emotional affect is highly connected with sexual interest, and as such it is possible that this additional factor may also play a role in any discovered interactions. Further exploration into how these factors may co-vary with pupil size is required to understand these interactions. Moreover, it is apparent that experiments should consider the influence of factors such as familiarity when investigating cognitive measures, as they may moderate observed relationships between cognitive measurement and factors such as sexual interest.

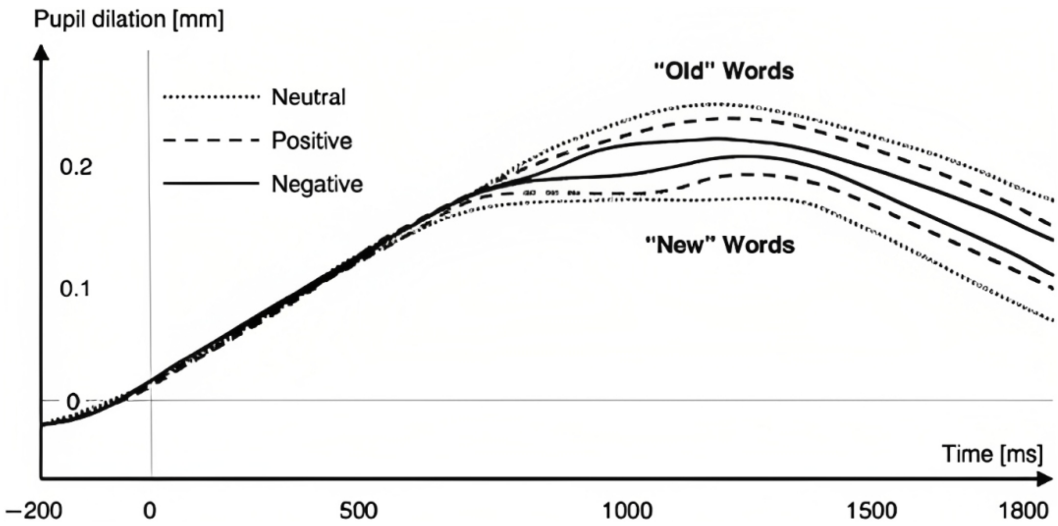

Figure 1.6. Graph detailing the mean pupil dilation curves for words with different emotional content (Taken from Võ et al. (2008).

1.5.2 FAMILIARITY AND ATTRACTIVENESS

Familiarity and attraction are inherently interlinked concepts, whereby familiarity influences subjective evaluations of attractiveness (Little et al., 2011;

Peskin & Newell, 2004; Thiruchselvam et al., 2016). The interaction between these two factors has been extensively explored using the mere exposure paradigm (Han et al., 2020; Moreland & Beach, 1992; Rhodes et al., 2005), whereby familiarity is directly manipulated by repetitive presentation of a stimulus. In these experiments, evidence suggests that just the act of repetitively viewing a face leads to these faces being viewed as increasingly attractive to participants. This is a clear caveat for experiments analysing attraction, where the potential influence of familiarity should be duly considered.

1.5.3 FAMILIARITY AND LIKING

Familiarity and liking are another pair of interconnected factors that may influence cognitive measures. These factors are inherently interlinked, as pre-existing familiarity (such as with a celebrity or famous figure), is often accompanied with a degree of liking for an individual. As such, it is pertinent to explore any possible interactions. Some evidence has suggested that increased likeability leads to increased perceived familiarity (Monin, 2003). However, the majority of research considers this relationship from the other direction, alluding to a possible bidirectional relationship. Evidence has suggested that by increasing familiarity through mere exposure, such as simply being presented with the same image multiple times, ratings of liking can be heightened (Brockner & Swap, 1976; Liao et al., 2011; Moreland & Zajonc, 1982). This has also been evidenced when considering increased liking for familiar public figures (Harrison, 1969). While mere exposure research often utilises laboratory conditions, with highly controlled stimuli, this effect has also been detected using more ecologically valid paradigms. Reis et al. (2011) investigated the familiarity-liking relationship through use of live interaction. In this experiment, participants met for the first time and interacted for varying time periods. Results showed that interpersonal attraction, and thus liking, depended on how much they had interacted. Therefore, compelling evidence suggests that familiarity produces liking.

This same trend of results has been discovered when familiarity is induced through the act of information sharing, such as answering pre-determined questions in the presence of another participant, rather than simple mere exposure (Sprecher et al., 2013; Ullrich et al., 2013). However, care should be taken when drawing conclusions from these results as some research has found the opposite trend of results (Norton et al., 2007, 2013). Finkel et al. (2015) propose that these conflicting results may be due to cognitive fluency, which refers to the ease in which we can understand information. Thereby, if the provided information is congruent with previous information then cognitive fluency is increased. This then leads to greater liking, a phenomenon also often occurring within the mere exposure paradigm. However, if the information is not easily integrated with prior knowledge, this cognitive fluency is decreased, reducing liking. Regardless of direction, it appears clear that familiarity can influence levels of liking. This highlights an important relationship that must be considered when investigating familiarity.

1.5.4 FAMILIARITY, ATTRACTIVENESS AND LIKING

It is also important to consider the potential interactions between these two factors, attractiveness and liking, when considering familiarity. As evidence has shown, relationships exist between these factors and familiarity in isolation, which may exert influence over cognitive measures. However, if these factors also interact, they may have a combined effect on familiarity. The perception of attractiveness can be affected by a number of different factors, with one such factor being liking (Kniffin

& Wilson, 2004). As such, it appears possible that increased liking for another may directly influence perception of attractiveness. However, particularly compelling is this link in the opposite direction. Many experiments have indicated that attractive individuals are evaluated and treated more positively than less attractive people, with evidence consistently showing that individuals rated as highly attractive also receive high ratings on other social dimensions such as liking and social skills (Eagly et al., 1991; Goldman & Lewis, 1977; Langlois et al., 2000). This link is evident in daily life, with attractive communicators liked more (Joseph, 1982) and the suggestion that attractive individuals may be more protected from work termination (Commisso & Finkelstein, 2012). In conclusion, it seems evident that attractiveness and liking are two interlinked concepts, which can exert influence over each other. Consequently, it is possible that these interlinked factors may also have a combined influence alongside feelings of familiarity. Throughout this thesis, the role of familiarity will be considered alongside sexual interest. As such, factors that can exert influence over this must be considered.

1.6 Structure of this thesis

The purpose of this thesis is to investigate whether cognitive measures of sexual interest are sensitive to differences in attractiveness within gender categories. The underlying factors that are explored throughout are initially investigated through self-report in the first experimental chapter. Familiarity, attractiveness and liking are explored and their potential links analysed. This precursory experiment considered the possible interactions between these concepts, creating a basis of understanding for further investigative study.

The second experimental chapter aimed to explore the key factors identified in chapter one, namely familiarity and attractiveness, within cognitive measures. Specifically, a series of experiments are presented which focus on within-category attraction. Cognitive tasks have seen huge success rates when considering attraction to a specific gender (Ciardha & Gormley, 2012; Ebsworth & Lalumière, 2012; Hess & Polt, 1960). However, limited research has considered whether such measures can indicate attraction within these categories. To this end, chapter two utilises a range of cognitive paradigms to investigate whether attractive and unattractive images of the same gender and identity produce differing cognitive responses. Furthermore, familiarity is explored throughout as a secondary factor to investigate whether any interactions exist with attractiveness in such measures.

The final experimental chapter specifically considers pupillary responses. This measure has been touted as a successful implicit measure of attraction (Attard-Johnson & Bindemann, 2017; Rieger et al., 2015; Rieger & Savin-Williams, 2012). However, these analyses again take place across gender groups, with little consideration within these categories. As such, this chapter explores the sensitivity of this response within gender. The second aim of this chapter is to explore the implicit nature of this measure. Observer presence is varied and participant attention directed to investigate whether pupil size can be manipulated or whether this response is impervious to control.

Chapter 2

Explicit Measures of Familiarity, Attractiveness and Liking

Experiment 1

Cognitive measures are frequently employed to analyse implicit attitudes and feelings, such as attraction. This concept has seen empirical investigation through cognitive measurement of sexual interest, with measures such as viewing time (Israel & Strassberg, 2009; Lippa, 2012; Quinsey et al., 1996), Stroop tasks (Ciardha & Gormley, 2012; Mannfolk et al., 2023) and pupillary responses (Attard-Johnson et al., 2021; Rieger & Savin-Williams, 2012; Watts et al., 2017) showing clear patterns congruent with sexual orientation. However, sexual interest is not the only factor known to affect cognitive measures, with factors such as familiarity and liking exerting influence over pupil size, too. For example, familiar stimuli have been found to elicit larger pupil dilation than novel stimuli (Kafkas & Montaldi, 2015; Naber et al., 2013; Weiss et al., 2016), through a phenomenon referred to as the pupil old/new effect (Võ et al., 2008). As factors such as sexual interest and familiarity can affect cognitive measures separately, it is plausible that they also may interact within such measures. This phenomenon is explored in later chapters in this thesis. However, before possible interaction effects can be investigated within *implicit* measures of sexual interest, it is important to understand how factors may interact separately when observers are asked *explicitly* to assess familiarity, attraction and liking.

Evidence suggests that when considered separately from implicit cognitive measures, familiarity and attraction frequently interact (Halberstadt et al., 2003; Little et al., 2014; Moreland & Beach, 1992). A common methodology employed to analyse this trend is the mere exposure paradigm, whereby participants are repeatedly shown the same stimuli to induce familiarity. In an experiment by Peskin and Newell (2004), participants were exposed to 16 faces six times each, and a further eight faces only once. As evidenced through attractiveness ratings, this simple increase in exposure led to higher levels of attraction. Therefore, it seems evident that familiarity and attraction may be interrelated concepts and, as such, this should be duly considered when investigating the interaction of these factors using implicit cognitive measures, too.

When considering the potential interactions present between sexual interest and familiarity, there are other factors that should be taken into account in order to avoid confounding effects. The concept of liking often occurs alongside attraction, in that we tend to like those we are attracted to (Commisso & Finkelstein, 2012; Joseph, 1982). Evidence for this suggestion can be found when observers evaluate face photographs, whereby increases in facial attractiveness lead to higher ratings of liking (Goldman & Lewis, 1977; Langlois et al., 2000). Similar results have been observed in real life environments, such as evaluations of politicians' faces in electoral settings (Efrain & Patterson, 1974). Due to the highly related nature of attraction and liking, these factors need to be effectively disentangled in order to discern which factor is being measured. This is important when considering the evaluation of these concepts using implicit cognitive measures. Here, precursors to liking, such as emotional affect, have been effectively measured through implicit cognitive measures such as Stroop tasks (Kappes & Bermeitinger, 2016; Wingenfeld et al., 2006) pupillary responses (Bradley et al., 2008; Janisse, 1973; Snowden et al., 2016). As such, potential confounds such as liking should be duly considered when drawing conclusions from cognitive measures of attraction.

While feelings of liking have been shown to influence attraction, these interactive effects may also extend to familiarity, the second factor of interest in this thesis. Extensive evidence suggests that increasing familiarity leads to higher levels of liking (Brockner & Swap, 1976; Moreland & Zajonc, 1982; Saegert et al., 1973; Sprecher et al., 2013). This is particularly evident in naturalistic environments, such as live interaction tasks. In an experiment where participants met for the first time during face-to-face interaction, with familiarity level controlled by conversation topic, higher levels of liking were recorded as participants became more familiar with one another (Reis et al., 2011). However, some evidence also indicates that this relationship is bidirectional, in that familiarity may *reduce* how much we like others (Norton et al., 2013; Sanbonmatsu et al., 2012). A series of experiments was conducted by Norton et al (2007), where participants were asked to fill out surveys indicating familiarity through how much knowledge they held about another person, and asked to rate them based on liking. The results showed that the more familiar participants were with someone, the less they liked them. Therefore, evidence suggests that the nature of this familiarity-liking link can be complex and sometimes contradictory, with further research suggesting that the direction of this interaction is moderated by relationship stage (Finkel et al., 2015). Regardless of the direction of this relationship, it is clear that familiarity and liking can interact. Therefore, this indicates that *attraction* and familiarity can also interact with feelings of liking. As such, this is a concept that should be considered when measuring these factors with implicit cognitive measures. However, before this can be considered, it is important to understand how these factors interact at a conscious, explicit level. As such, this study will investigate familiarity, attraction and liking through a self-report survey, serving to inform later thesis chapters where implicit cognitive measures will be employed. Furthermore, two types of stimuli (names and faces) will be presented to explore their utility in eliciting feelings of familiarity, attraction and liking.

Facial images (Little et al., 2014; Monin, 2003; Peskin & Newell, 2004) are often presented when investigating these factors, with limited research utilising name stimuli (Erwin, 1993; Hensley & Spencer, 1985). Both names and faces provide clues for identity, with names capable of activating *internal* representations of a person's appearance whereas faces also provide *external* representations of identity. Cognitive models of person recognition also specify that access to cognitive representations of a person's name and face are structurally linked, whereby activation of one type of stimulus can facilitate processing of the other (see, e.g., Burton et al., 1990). Therefore, if names produce the same relationship between feelings of familiarity, attraction and liking as faces do, then this would indicate that evaluations of faces map onto the same internal person representations as names. On the other hand, if dissociable relationships between these factors are found for faces and names, then this would indicate that observers evaluate familiarity, attraction and liking differently depending on whether this information has to be perceived from an external stimulus or retrieved from internal representations.

Method

Participants

Fifty-two participants, (26 male, 26 female) with a mean age of 20.31 years (*SD* = 5.20, Range = 18 to 54), participated in this experiment. All participants were aged 18+ and had normal-to-corrected vision. Three participants were excluded due to non-completion of the experiment, resulting in a final sample of 49 participants.

Materials

The stimuli consisted of facial images of 50 celebrities (25 female and 25 male), and their names. The images were obtained from screen shots of the Graham Norton show and feature each celebrity portraying a relatively neutral expression (see Figure 2.1 for example images). Each image was cropped to a height and width of 7.7 cm. Faces were angled slightly away from the camera due to the camera position used in the show. This method allowed for more naturalistic images compared to previously adopted methods such as the use of passport style photographs. Image background was comparable across stimuli due to collection from the same source. The name stimuli featured only the names of each celebrity. In most cases this consisted of both their first and last name, but in cases where the celebrity is best known by a single name, one name was presented to avoid potential confusion (e.g. RuPaul).

Figure 2.1. An array featuring examples of the face stimuli employed in this experiment.

At the beginning of each trial in both blocks, the stimuli were presented alongside the question *'How much do you know about this celebrity'*, with a 7-point scale from *'1 – Nothing at all'* to *'7 – A lot'*. This acted as a measure of Familiarity and was used to determine the questions that were presented subsequently. Within the name block, in the case of complete unfamiliarity (an indication of 1), the trial would immediately move on as no further information could be garnered. However, if familiar with the celebrity (i.e., a score larger than 1), participants were presented with six questions, designed to assess Familiarity, Attractiveness and Liking. These questions are summarised in Table 2.1 and were aimed at understanding how these factors may interact within an explicit self-report task.

A similar format was implemented for the image block, with slight differences, see Table 2.2. In this block, if a celebrity was unfamiliar, participants received only the three questions pertaining to Attractiveness and Liking, as a face image is adequate for these judgements. However, if familiar, all six questions regarding Familiarity, Attractiveness and Liking were presented.

Procedure

This experiment was provided online through Qualtrics and consisted of two separate blocks. In the first block, participants were provided with the name of a celebrity and a 7-point Likert scale pertaining to how familiar they were with the presented celebrity. Subsequent questions surrounding Familiarity, Attractiveness and Liking were then provided depending on the initial familiarity judgement. In the second block, participants were provided with the facial image and name of one of the celebrities, alongside the same 7-point scale to assess Familiarity. Again, the presentation of subsequent questions depended on the initial familiarity score. The experiment was completed once all ratings were provided for each celebrity.

Table 2.1

All possible questions presented within the name block

Questions	Answer Scale (Min. to Max.)
1) How much do you know about this celebrity?	1- Nothing at all 7- A lot
2) What is this celebrity's profession? (Please select as many as apply).	1- Unknown 2- Comedian 3- Actress/ Actor 4- Singer 5- Sport Personality 6- TV Personality
3) Do you know what this celebrity looks like?	1- Definitely not 7- Definitely yes
4) How 'good looking' would you consider this celebrity to be, irrespective of whether you are personally attracted to them?	1- Not at all 7- Extremely
5) From your personal perspective, how sexually attractive do you think this celebrity is based on their appearance?	1- Not at all 7- Extremely
6) From your personal perspective, how sexually attractive do you think this celebrity is based on their personality?	1- Not at all 7- Extremely
7) How much do you like this celebrity based on their personality?	1- Not at all 7- Extremely

Table 2.2

All possible questions presented within the face block

Questions	Answer Scale
8) How much do you know about this celebrity?	1- Nothing at all 7- A lot
9) How well do you think this image captures the typical appearance of this celebrity?	1- Not at all 7- Extremely
10) How 'good looking' would you consider this celebrity to be, irrespective of whether you are personally attracted to them?	1- Not at all 7- Extremely
11) From your personal perspective, how sexually attractive do you think this celebrity is based on their appearance?	1- Not at all 7- Extremely
12) From your personal perspective, how sexually attractive do you think this celebrity is based on their personality?	1- Not at all 7- Extremely
13) How much do you like this celebrity based on their personality?	1- Not at all 7- Extremely
14) In reference to level of physical attractiveness, how similar to you is the celebrity?	1- Not at all 7- Extremely
15) Overall how similar to you is the celebrity?	1- Not at all 7- Extremely

Results

In the name condition, any Familiarity rating above 1 was initially taken to indicate familiarity, prompting the presentation of all further questions in this block. However, for the purpose of data analysis, ratings above four were considered to indicate familiarity, due to this being the midpoint of the scale. As such, results will reflect familiarity apart from selected instances, which will be specified, in which results below this familiarity score are explored. This allows for comparisons across the name and faces block as, due to the nature of word stimuli, Attractiveness and Liking ratings could only be collected for familiar celebrities. All subsequent correlational data has been amalgamated and is provided in Table 2.3. In order to undertake these analyses, the scores for each question was calculated across all participants for each celebrity. Analysis was then conducted, not on specific celebrities, but correlations were computed for each question across all average data points across celebrities. Furthermore, post-hoc power analysis was completed on G*Power to test the statistical power of this experiment achieved. With a sample size of 49, and r = .83, the achieved power equals 1.00.

Table 2.3

P-values of correlations between all items within the names and faces blocks when considering those who indicated a familiarity score above four.

Concept	Items	Names (*p*-value)	Faces (*p*-value)
Attractiveness and Liking	Good looks – Attraction to personality	**< .001*** **	**0.001** **
	Good looks – Liking for personality	.11	.32
	Attraction to appearance – Attraction to personality	**< .001*** **	**< .001*** **
	Attraction to appearance – Liking for personality	**< .001*** **	**.004** **
Familiarity and Attractiveness	Knows – Attraction to appearance	**.002** **	.24
	Knows – Good looks	.19	.07
Familiarity and Liking	Knows – Attraction to personality	**< .001*** **	.65
	Knows – Liking for Personality	**< .001*** **	.28

Note. *** indicates $p < .001$, and ** indicates $p < .01$.

Name Stimuli

Attractiveness and Liking

Firstly, Pearson's correlations were calculated to ensure that both items within Attractiveness and Liking were accessing their corresponding concepts. These correlations show that the items within Attractiveness (items 4 and 5), $r(47) = .83$, $p < .001$, and Liking (items 6 and 7), r(47) = .89, $p < .001$, were strongly and positively correlated. This confirms that both items within each concept of Attractiveness and Liking were measuring the same factor.

The relationships between Attractiveness and Liking were then explored by calculating Pearson's correlations between the two items within each of these concepts (see in Figure 2.2). A positive correlation between Attractiveness and Liking was found when considering good looks and attraction to personality, $r(47) = .50$, $p < .001$, attraction to appearance and attraction to personality, $r(47) = .74$, $p < .001$, and attraction to appearance and liking for personality, $r(47) = .55$, $p < .001$. This indicates that for celebrities whose appearance was accessed from their names, the more attractive they were considered to be, the more they were liked. However, while attraction to appearance correlated with liking for personality, no correlation was identified between the items of good looks and liking for personality, $r(47) = .23$, $p = .11$. This indicates that people who are judged to be good looking are not necessarily the same people whose appearance participants might be attracted to, and it is physical attraction rather than good looks *per* se that relates to Liking.

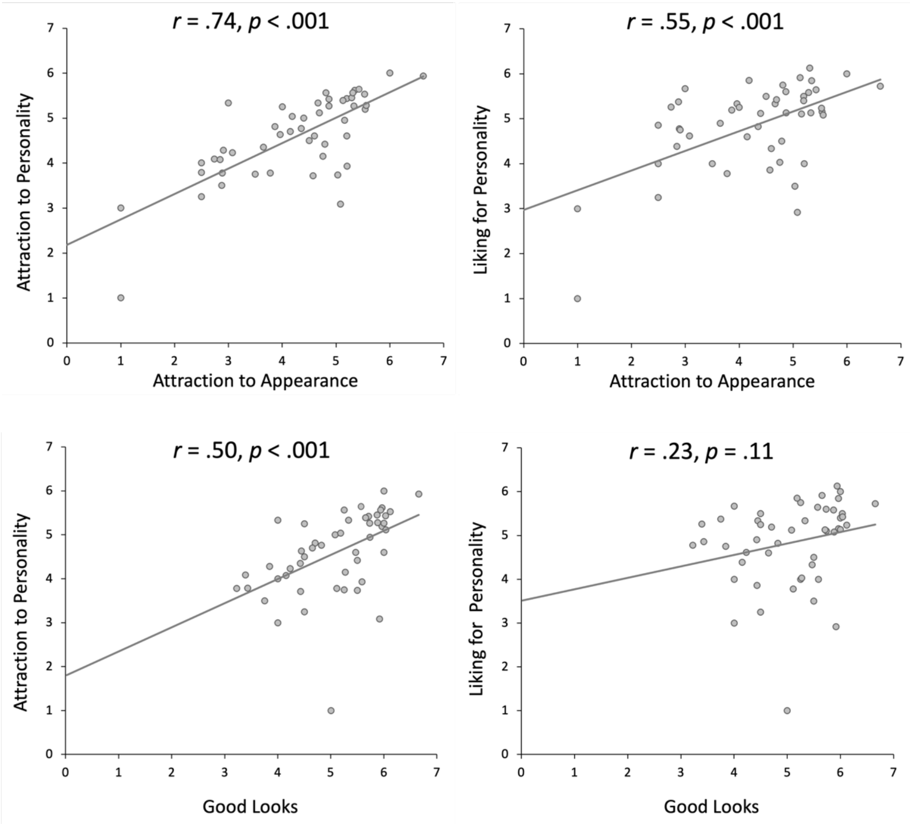

Figure 2.2. Scatterplot array featuring correlations between each of the two items indicative of Attractiveness and Liking, containing a regression line.

Familiarity

Correlations were then calculated between Familiarity and the other factors (see Figure 2.3). This revealed a positive correlation between Familiarity and one of the two scores of Attractiveness: attraction to appearance, $r(48) = .44$, $p = .002$. This showed that the more familiar a participant was with a celebrity, the more attractive they considered them to be. However, no correlation was identified for the Familiarity and the Attractiveness item of good looks, $r(49) = .19$, $p = .19$. When considering the relationship between Familiarity and Liking, the attractiveness of a personality, $r(48) = .52$, $p <.001$, and liking for a personality, $r(48) = .48$, $p <.001$, correlated with Familiarity. This indicates that the more familiar a celebrity is, the more likeable they are also thought to be.

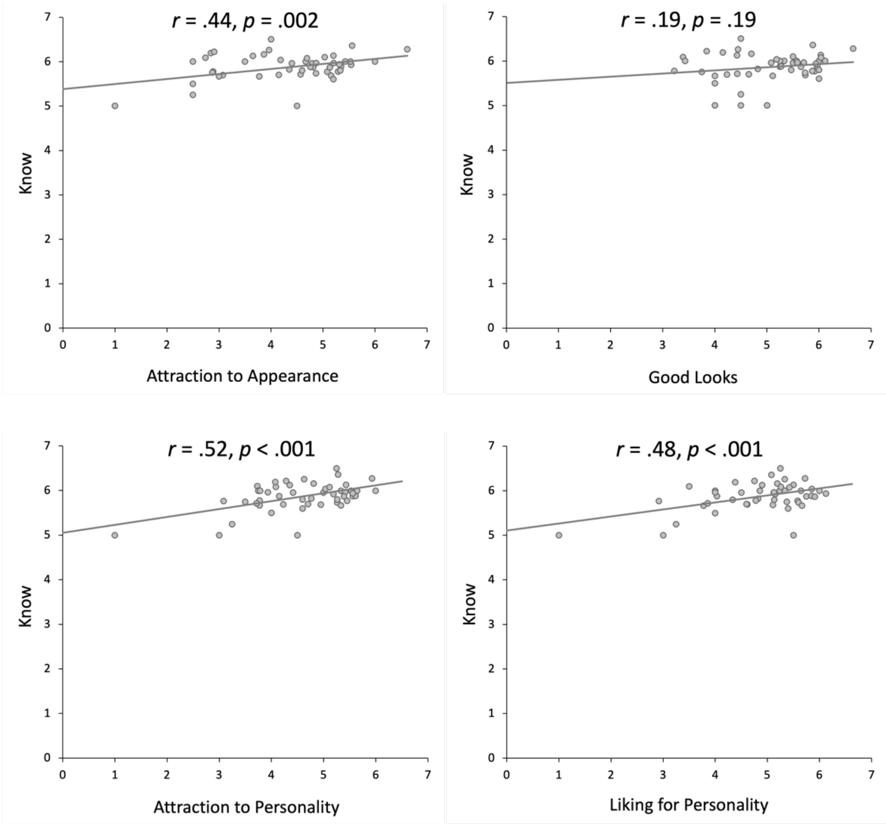

Figure 2.3. An array of scatterplots, featuring regression lines, reflecting participant ratings of Familiarity alongside Attractiveness and Liking.

Face Stimuli

Attractiveness and Liking

Correlations were again conducted to ensure that the items within Attractiveness and Liking were accessing the same concepts, with results showing that both Attractiveness (items 12 and 13), $r(48) = .72$, $p < .001$ and Liking (items 14 and 15) were highly correlated, $r(48) = .80$, $p < .001$. Next, Pearson's correlations were computed for participants who indicated Familiarity with the celebrity stimuli. As can be seen in Figure 2.4, positive correlations were found for good looks and attraction to personality, $r(48) = .44$, $p = .001$, and for attraction to appearance and attraction to personality, $r(48) = .70$, $p < .001$, and attraction to appearance and liking for personality, $r(48) = .40$, $p = .004$. This suggests that the more attractive participants considered a celebrity to be, the more they liked them. However, as with name stimuli, no correlation was discovered between the items of good looks and liking for personality, $r(48) = .14$, $p = .32$, again indicating that people who are judged to be good looking are not necessarily the same people whose appearance participants might be attracted to, and it is physical attraction rather than good looks that relates to Liking.

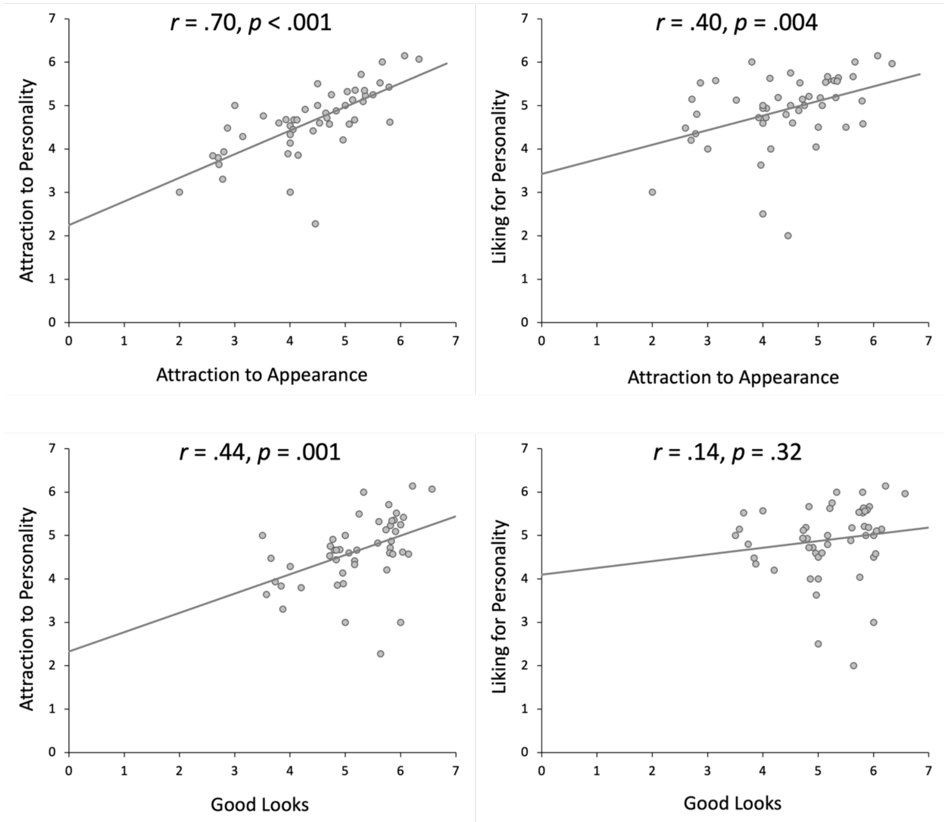

Figure 2.4. Scatterplots representing the correlations, with regression line, between Attractiveness and Liking in the face condition.

Further analysis was carried out to investigate how Attractiveness and Liking are associated in the viewing of less familiar facial stimuli. For this analysis, only participants who indicated a score below 4 on the 7-point scale of Familiarity were considered. However, there were no correlations found between any items, all *ps* > .13 (see Figure 2.5). This indicates that, in contrast to familiar faces, Liking and

Attractiveness are dissociable for unfamiliar faces. This seems logical given that participants have no insight into the personalities of people they do not know, and demonstrates that such information is not extrapolated falsely from impressions of Attractiveness.

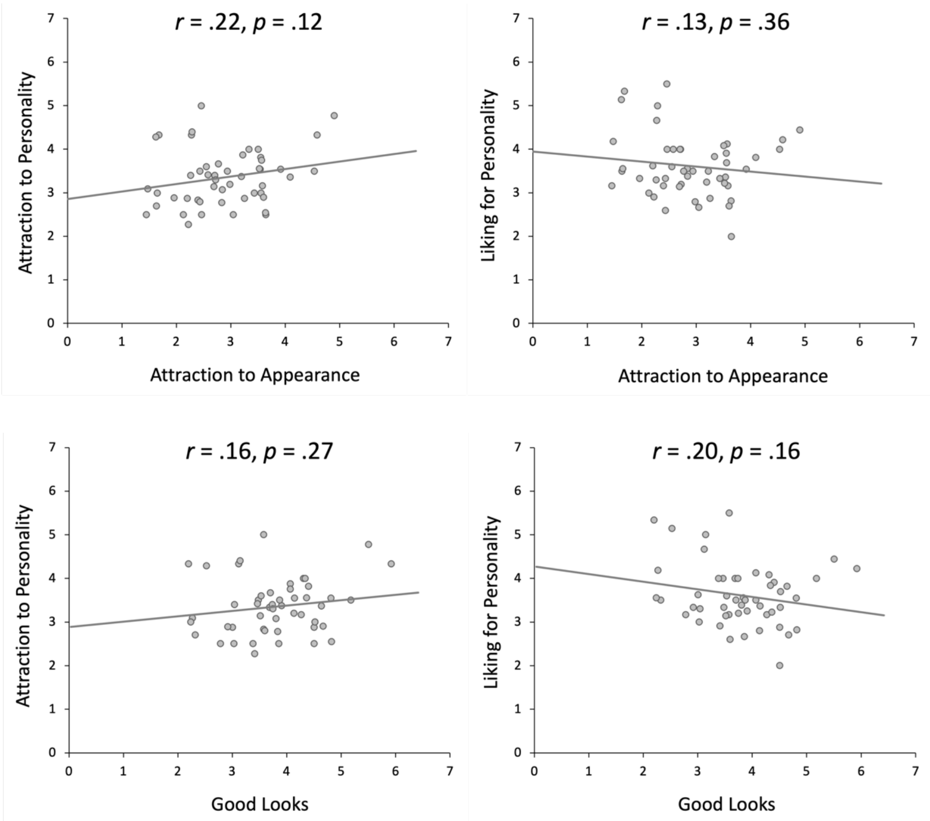

Figure 2.5. Scatterplots representing the correlations between the items of Attractiveness and Liking when participants were unfamiliar with the celebrity face stimuli.

Familiarity

Finally, Familiarity was considered alongside Attractiveness and Liking. As can be seen in Figure 2.6, Pearson's correlations indicated no significant correlations between Familiarity and the Attractiveness measures of good looks, $r(48) = .26$, $p = .07$, and attraction to appearance, $r(48) = .17$, $p = .23$. Similarly, there were no reliable relationships between Familiarity and the attraction to personality, $r(48) = .07$, $p = .64$, and liking for personality, $r(48) = .16$, $p = .27$. This suggests that when presented with a facial image, and thus when judgements are not reliant on memory, Familiarity, Attractiveness and Liking act as independent concepts.

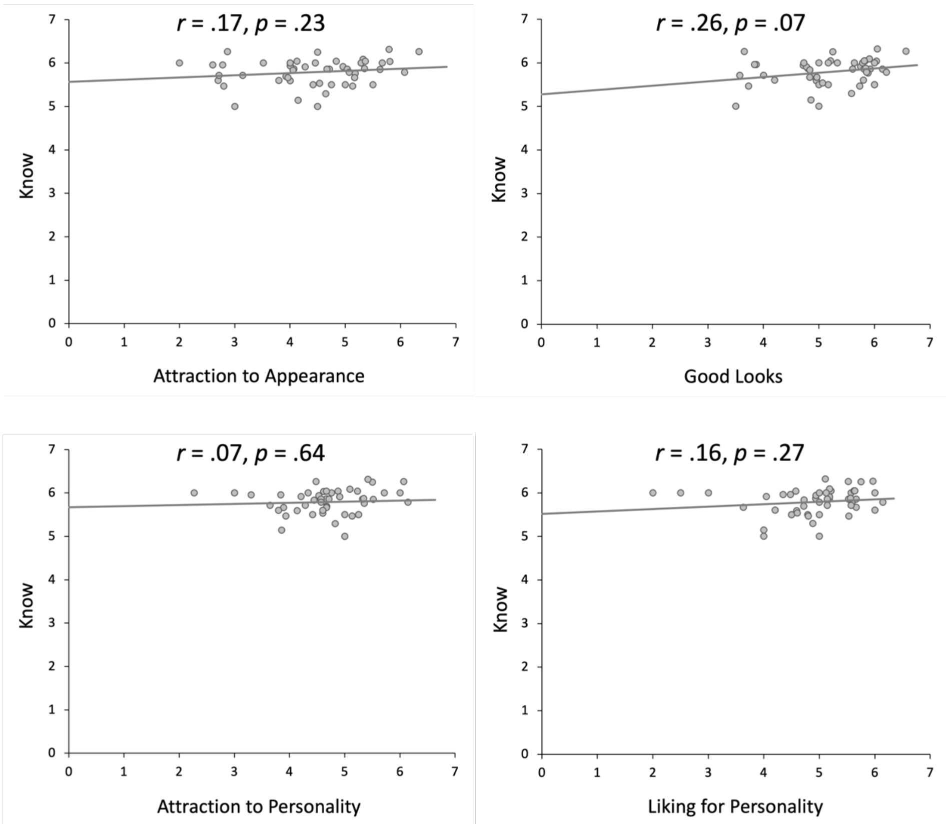

Figure 2.6. Array of scatterplots for the correlations between items indicative of Familiarity, Attractiveness and Liking when presented with facial images.

Discussion

This experiment investigated interactions between Familiarity, Attractiveness and Liking based on self-report ratings, to examine the relationships between these factors in the evaluation of people. This study also explored whether the relationship of these factors depends on whether observers are viewing faces or have to access internal representations of a person's appearance via their name. These results were collected to provide a basis for subsequent experiments that examine these factors with implicit cognitive measures.

Firstly, we consider judgements triggered by name information to discover if stimuli type affected how ratings were performed. Correlations looking at the relationships between the two items indicative of each Attractiveness and Liking suggested that, overall, celebrities who were considered as more attractive, were generally liked more. This finding was expected, due to its congruence with the wider literature (Joseph, 1982; Kniffin & Wilson, 2004; Langlois et al., 2000). Only one of these relationships, between the items of good looks and liking for personality, did not correlate, explanations for which will be explored later in this chapter. Despite this, both of the items within Attractiveness and Liking demonstrated a highly positive relationship, suggesting a correspondence in measured factors.

When expanding this analysis to include Familiarity, a clear positive relationship was identified between Familiarity and Liking. Only one item within Familiarity was used in this analysis (know), as this item effectively encapsulates the concept of familiarity. Results indicated that generally the more familiar the participant was with a celebrity, the more they liked them, further corroborating findings from wider literature (Harrison, 1969; Sprecher et al., 2013). The interaction between Attractiveness and Familiarity showed a more complex relationship as only one of the Attractiveness items (attraction to appearance) correlated with the Familiarity item, with no correlation identified for the good looks item. The absence of a correlation between good looks and either item of Familiarity and Liking hints at the type of attraction that may exist within the relationships between these factors. Despite the high correlation between the two Attractiveness items, they seem to access different facets of this factor. Good looks refers to a more objective judgement of someone's facial appearance, while attraction to appearance reflects the more subjective judgment of how attracted one person is to another individual. As such, it seems that feelings of attraction, rather than the objective attractiveness of others, interacts with Familiarity and Liking.

When considering facial stimuli, when participants were familiar with the celebrities, strong positive correlations were identified between the Attractiveness item of attraction to appearance and both items of Liking (attractiveness to personality and liking of personality). Once again, no relationship was uncovered between Liking and the Attractiveness item good looks, further consolidating the idea that Liking interacts with subjective attraction, rather than simply how good looking an individual generally is perceived to be.

However, the link between Attractiveness and Liking for less well-known celebrities is also of interest in this block. This analysis was impossible for name stimuli, as when presented with a name, Attractiveness and Liking evaluations can only be made for those we already have a mental representation of. However, these concepts can be inferred from facial images. Results for unfamiliar stimuli showed that no correlations were identified between any of the items of Attractiveness or Liking.

When expanding analysis to consider ratings of Familiarity alongside Attractiveness and Liking, no significant correlations were identified. This is unlike the findings identified within the name condition whereby positive correlations were found amongst all factors. This may be due to variance in familiarity cues between conditions. When provided with a name, familiarity is treated as more of a binary yes or no indication, as there is little information to cue familiarity beyond a simple, do I know someone with this name, decision. Within this more simplistic categorisation of Familiarity, Attractiveness and Liking, evaluations are confounded by Familiarity, in that we perceive familiar people to be more likeable because they are attractive, or vice versa. However, when provided with a facial image, familiarity is treated as more of a scale. These images provide more familiarity cues, as observers may be somewhat familiar with the celebrity, such as recognising them from a film or a show, without knowing them well enough to be able to identify their name. In these cases, there is higher variance in Attractiveness and Liking ratings at each level of this Familiarity scale, and these are not related. As such, the results suggest that Familiarity, Attractiveness and Liking are treated as separate and independent concepts when facial images are provided due to the increased presence of familiarity cues. It should also be noted that there is a restricted range of correlations when considering the factor of Familiarity, due to the exclusion of data falling below a rating of 4. While this was conducted to categorise familiar and unfamiliar others, any potential correlations with Attractiveness and Liking may have been limited as a result.

The aim of this experiment was to analyse how Attractiveness, Familiarity and Liking interact to form a basis of understanding for future experiments evaluating these factors using implicit cognitive measurements. Within the name condition, the results generally mimicked those found in the wider literature (Commisso & Finkelstein, 2012; Harrison, 1969; Monin, 2003), where these factors are often related. However, one of the items of Attractiveness (good looks) did not follow this same pattern of results. This has implications for our understanding of exactly how, and when, these factors interact. The lack of relationship present due to the good looks item suggests that Attractiveness only interacts with the other factors when considering personal attraction to another, and not just the objective attractiveness of faces. As such, when considering how these factors may interact within cognitive tasks, the nature of the factors included should be considered. Consequently, when assessing attraction in subsequent thesis chapters, personal attraction will be considered rather than objective judgements in order to facilitate potential interactions between factors.

Furthermore, the type of stimuli had a considerable impact on this experiment. While the name block showed relationships present between Familiarity, Attractiveness and Liking, this was not identified in the face condition when considering Familiarity. Neither Liking or Attractiveness correlated with this factor, suggesting that Familiarity interacts differently when presented with a face than when recalling an identity cued by a name. While the potential differences between names stimuli and facial images when evaluating feelings of Familiarity, Attractiveness and Liking have not yet been widely explored, this study shows that stimuli type can significantly affect the ratings of these attributes, particularly when considering Familiarity. The type of stimuli included in this experiment should also be evaluated. The stimuli consisted of the names and faces of various celebrities and thus the generalisability of these findings to everyday others should be considered. While this cannot be addressed in the current experiment, a further study analysing these groups alongside each other, while implementing the current design, could effectively investigate these generalisability concerns.

In conclusion, when considering facial images and names of celebrities, ratings of Familiarity, Attractiveness and Liking often correlate. Generally, the more we like someone, the more attractive we find them and vice versa. Furthermore, the more familiar we are with another person, the more attractive and likeable they are usually considered to be. This demonstrates that evaluations of Attractiveness, Liking and Familiarity are intertwined. This raises the question of whether widely used implicit cognitive measures of attractiveness can be affected by the interplay of these factors. This question will be examined in Chapter 3.

Chapter 3

Implicit Measures of Sexual Interest

Introduction

Human beings are inclined to evaluate the attractiveness of others, with judgements occurring automatically (Ritchie et al., 2017) and within a fraction of a second (Locher et al., 1993; South Palomares & Young, 2018; Willis & Todorov, 2006). One facet of these evaluations pertains to sexual *orientation*, which acts as a lens through which *attraction* to another person can be determined. As such, sexual orientation has been investigated extensively, traditionally in the form of self-report (Dombert et al., 2016; Hoon et al., 1976; Howell et al., 1987), such as in Chapter 2 of this thesis which explored attractiveness alongside the secondary factors of familiarity and liking, via a self-report questionnaire. However, due to the sensitive and social nature of sexual orientation, the potential manipulability of this response has led to the development and implementation of more implicit cognitive measures.

These cognitive measures of sexual orientation are often based on the measurement of latency and attention. A classic example of this is the Viewing Time paradigm (Rosenzweig, 1942). Viewing time tasks utilise the assumption that images congruent with sexual orientation create an attention-driven delay, resulting in longer viewing times (Gress et al., 2013). These findings have been replicated, albeit to varying degrees, across experiments utilising an extensive range of stimuli including nude (Harris et al., 1996; Quinsey et al., 1996), partially clothed (Israel & Strassberg, 2009; Rönspies et al., 2015), and fully clothed imagery (Schmidt et al., 2017). The general robustness of this measure for detecting attraction to specific person categories has led to implementation within forensic populations (Gray et al., 2015; Gress, 2005; Schmidt et al., 2017).

A range of other paradigms have been developed to provide implicit measurements of attraction, such as the Pictorial Stroop. This task is based on the original Stroop test (Stroop, 1935) and involves the presentation of images with coloured tints. In this paradigm, participants are typically slower in identifying image hue for stimuli of the attractive gender category (Ciardha & Gormley, 2012; Mannfolk et al., 2023). These pictorial Stroop tasks allow for direct presentation of sexually attractive individuals, making them advantageous over word-based Stroop tasks, for use in research where stimuli pose, clothing, gender, or identity are under direct observation (Ciardha & Gormley, 2009).

From viewing time measures (Harris et al., 1996; Quinsey et al., 1996) to Stroop tasks (Ciardha & Gormley, 2012; Price & Karl Hanson, 2007), results show that heterosexual individuals view other-sex stimuli for longer than same-sex stimuli. Generally, however, these tasks have demonstrated success in reflecting sexual orientation of male participants, with mixed results reported for other-gendered individuals (Bailey, 2009; Lippa, 2012; Quinsey et al., 1996).

Throughout these experiments, attractiveness is often treated as a simple category, whereby stimuli are divided into images corresponding to observers' sexual orientation (male vs. female). However, some experiments have sought to gain understanding of the sensitivity of these measures beyond their category-based natures. One experiment discovered that male and female participants viewed more attractive models for longer than less attractive models (Lippa, 2012). For men this was specific to female models, whilst women followed this pattern regardless of sex category. However, there is also opposing evidence to suggest that viewing time measures are less sensitive to these stimulus-specific findings. When viewing nearly-nude magazine models, Israel and Strassberg (2009) found that viewing time and attraction ratings did not correlate, despite longer viewing times for opposite sex images. Thus, whereas viewing time can consistently measure broad-level attractive impressions of sex category, it is unresolved whether similar effects persist at a more fine-grained level within groups.

Resolving this issue has theoretical advantages and disadvantages, both as a measure and as a human phenomenon. When considering attractiveness at category level, there are merits within forensic applications. For example, when investigating sex offenders with child victims, it is pertinent to identify age-related sexual interest, as this leads to increased recidivism rates (Barbaree & Marshall, 1988; Hanson & Bussière, 1998; Seto et al., 2004). However, in these cases, attractiveness towards individual stimuli has not been assessed. Such within-category distinctions may be insightful when considering factors that may interfere or interact with feelings of attractiveness, such as familiarity with a person (Han et al., 2020; Moreland & Beach, 1992; Peskin & Newell, 2004). As such, it is important to uncover what level of attractiveness cognitive measures can reflect. Currently, it remains unresolved whether category attractiveness defines behaviour irrespective of other variation between individuals within categories, or whether category responses are driven by such differences.

When considering cognitive measures of attraction, stimuli often vary in identity. This can lead to potential confounds through individual preferences in attraction to stimuli models. Moreover, differences in attraction between people are also accompanied by within-person variability, a term referring to the inevitable fluctuations in appearance experienced by the same individual. The presence of such within-person variation has been identified through sorting tasks, where appearance deviates so greatly within individuals that photos of the same identity are often classified as representing different people (Jenkins et al., 2011). This within-person variation is so extensive that in order to accurately learn a face, exposure to multiple instances of identity is required (Andrews et al., 2015; Murphy et al., 2015; Ritchie & Burton, 2017). As a consequence of such variation, attractiveness of an individual also differs across images, to the point where this can vary more *within* an identity than *between* individuals (Jenkins et al., 2011). This concept has rarely been utilised in the study of sexual attraction, as research has focused on response to unfamiliar others, making it difficult to determine whether an attractive *person* or an attractive *image* is being viewed. However, by presenting participants with two images of the same identity that vary in attractiveness, confounds introduced through individual differences can be avoided.

In the current series of experiments, a range of paradigms and stimuli will be employed to investigate differences in attractiveness across and within the sex category of person stimuli. This will allow for a more robust exploration into the level of attractiveness that cognitive measures can assess. In Experiments 2 and 3, viewing time tasks will be utilised to analyse attraction at both the between- and within-category levels of sexual attraction, whilst Experiment 4 collapses this design and only considers within-category levels of attraction. Akin to the aforementioned literature, we expect to find longer viewing times for images congruent with sexual orientation, specifically in heterosexual male participants. However, it remains unknown whether these viewing-time effects will be present across within-category variation in attractiveness. In Experiment 5, a Pictorial Stroop task will then be implemented to replicate and extend potential within-categorical findings. Across tasks, a range of stimuli will be employed to analyse the validity of cognitive measures as a response of between- and within-category attractiveness. Such stimuli will include full body images, in casual clothing and swimwear, as well as facial images. These experiments will lead to further understanding as to whether cognitive measures are tapping into attractiveness as a categorical concept, whereby people are deemed attractive based on the sex group they belong to (such as male or female) or whether such measures are sensitive to higher levels of distinction (i.e. interest in a specific-other).

Finally, as a further point of interest, familiarity with stimulus identity will also be considered throughout this study. The potential interactions between attractiveness and familiarity were investigated in Chapter 2 of this thesis, with correlations identified for name but not face stimuli. However, generally evidence has shown that familiarity can affect and interact with attractiveness through increased exposure, as repeated observation of a person or image leads to higher attractiveness ratings (Han et al., 2020; Peskin & Newell, 2004; Thiruchselvam et al., 2016). However, it remains unknown how familiarity as a factor may influence feelings of attractiveness within sex categories, as well as between categories. Thus, in the following experiments familiarity will be investigated through use of repeatedly presented known and unknown identities. This will allow for more robust investigation into the levels of sensitivity within cognitive measures, as well as further understanding into the potential interactions within these factors.

Experiment 2

Viewing time tasks are often implemented to explore different facets of attractiveness, such as sexual orientation (Ebsworth & Lalumière, 2012; Locher et al., 1993). These tasks are based on the notion that the longer a participant chooses to view a stimulus, the more attractive they deem it to be (Rosenzweig, 1942). In these experiments, analysis is often constrained to comparing categories of stimuli, such as men and women, to gain insight into category-based attractiveness. However, there has been limited investigation into the sensitivity of this measure *within* these sex groups, and with conflicting results (Israel & Strassberg, 2009; Lippa, 2012).

This experiment will first focus on the more widely identified *between*-sex group differences in attraction. This will further our understanding of whether implicit measures can assess differences in attraction, providing a strong basis for focus on *within*-sex group differences later in this chapter. To investigate this, images of male and female soap opera actors will be presented to participants. This experiment will additionally explore the role of familiarity in viewing time responses by featuring both familiar and unfamiliar identities. Evidence has shown that the more familiar a face, the more attractive it is considered to be (Han et al., 2020; Moreland & Beach, 1992; Rhodes et al., 2005). However, it remains unknown as to whether familiarity with a stimulus model will exert effects over higher-level distinctions of attractiveness, and whether these effects can be observed through implicit cognitive measures.

Method

Participants

This experiment included 80 heterosexual participants, comprising of 40 men (M_{age} = 40.6 years, *SD* = 15.02, Range = 20 to 75) and 40 women (*M* = 37.97, *SD* = 12.29, Range = 21 to 64). Two male participants were excluded from analysis due to aborting the experiment before completion. All participants resided in the UK at time of participation, identified as heterosexual according to the Kinsey scale (Kinsey et al., 1948), with 95% of males and 90% of females reporting UK nationality.

Materials

The experimental stimuli consisted of 20 British and 20 German TV Soap Opera stars, see Figure 3.1 for example stimuli. The British celebrities (10 male, 10 female) were selected from a list of former and current *EastEnders* cast members, while the German celebrities (10 male, 10 female) were selected from the popular German Soap *Gute Zeiten, schlechte Zeiten.*

Figure 3.1. Example stimuli for Experiment 2. Familiar soap opera celebrities are featured in the top row, and unfamiliar in the bottom row.

These celebrities were chosen as they represent a population which is nationally famous, in that they are likely to be familiar to people living in the same country as the show is aired, but unfamiliar to all others. The use of soap stars allows for comparability across the superficial elements of the stimuli such as pose, lighting and image quality. This allows for more equitable stimuli across familiarity conditions than can be obtained through use of alternative populations, such as highly familiar A-list celebrities and unknown individuals.

Stimuli were collected using a google image search for the name of each celebrity followed by the phrase 'facial image'. The first high quality, frontal view image of each celebrity was chosen. Each image was then cropped from the background and resized to a height of 600 pixels, with a pixel per inch resolution of 72. The remaining facial image was then superimposed onto a plain grey background measuring a height of 27.09 cm and a width of 36.12 cm.

Procedure

This experiment was programmed in PsychoPy (Peirce et al., 2019) and implemented online via Prolific through Pavlovia. Participants were first presented with instructions asking them to rate the following images based on how attractive they found them to be. This wording was chosen to highlight the role of personal attraction in this decision-making process due to findings from Chapter 1 of this thesis indicating that general and personal attractiveness judgements lead to different interactions with additional factors, such as familiarity. A key press of '1' was considered 'unattractive' with '5' constituting 'highly attractive'. Instructions stated to respond as quickly and accurately as possible. Each stimulus appeared in a randomised order and remained on display until a response had been made. This was then repeated a further three times, for a total of four blocks, with a break included in the middle. A last block was then presented for a familiarity check, with the images each being shown a final time. This time participants were asked to rate how familiar they were with each individual model, from ‘1 – Unfamiliar’ to ‘5 - Highly Familiar’.

Results

All analyses were conducted with a familiarity check applied. This process involved comparing participants self-reported familiarity scores with the predetermined stimulus familiarity category, i.e. UK versus German celebrities. For the familiar UK category, data was excluded from analysis for identities categorised as unfamiliar, indicated by a score lower than ‘*3*’. For the unfamiliar German category, familiar identities were excluded, represented by a score higher than ‘*3*’. This ensured that both of these categories were representative of participant familiarity and not simply assumed based on nationality. This resulted in the exclusion of 47.70% of trials featuring familiar celebrities and 41.18% of those presenting unfamiliar celebrities. Post-hoc power analysis indicated that this experiment had limited statistical power of .25. In order to account for outliers in the data that may skew results, any viewing time over 6 seconds was excluded from analysis. Across the literature in this field there has been a wide range of exclusion criteria implemented, with outliers classified within a range of 5-15 seconds (Ebsworth & Lalumière, 2012; Lippa, 2012; Xu et al., 2017). A criteria of 6 seconds was applied as it fell within this range and provided a stringent value, allowing for higher confidence in any effects identified based on viewing time differences.

Male Participants

Ratings

Analysis was broken down by participant gender to investigate how results differ across these categories. Figure 3.2 represents the mean attractiveness ratings per stimulus category for male participants. A 2 (Sex: same, opposite) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA showed a main effect of Sex, $F(1,21) = 33.00$, $p < .001$, $\eta^2 = .31$, whereby people of the opposite sex were rated as more attractive than people of the same sex. There was no main effect of Familiarity found, $F(1,21) = 2.69$, $p = .12$, $\eta^2 = .02$, or an interaction between Sex and Familiarity identified, $F(1,21) = 0.54$, $p = .47$, $\eta^2 = .001$.

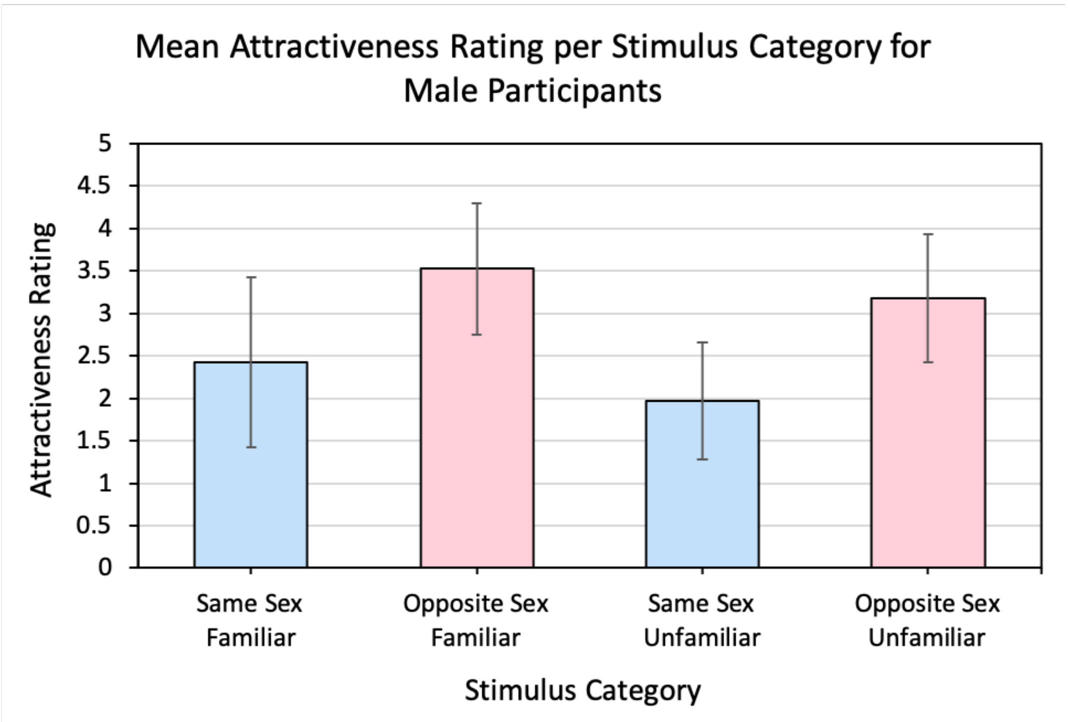

Figure 3.2. Bar graph showing average attractiveness ratings per stimulus category within male participants. Standard deviation is represented through error bars.

Viewing Time

A 2 (Sex: same, opposite) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA indicated no main effect of Sex, $F(1,21) = 0.16$, $p = .70$, $\eta^2 = .01$, or Familiarity, $F(1,21) = 0.00$, $p = 1.00$, $\eta^2 = .00$. There was no interaction found between these factors, $F(1,21) = 0.37$, $p = .55$, $\eta^2 = .02$. See Figure 3.3 for mean values.

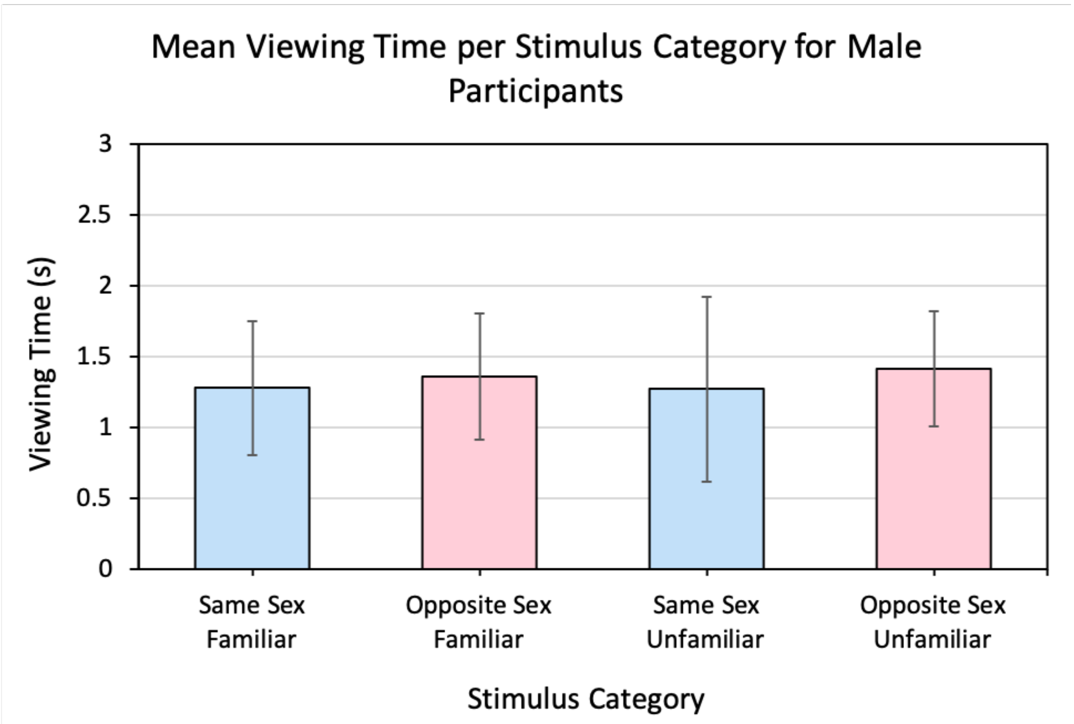

Figure 3.3. Bar graph depicting mean viewing times per stimulus category within male participants. Error bars show standard deviation.

Female Participants

Ratings

The analysis was then split to only consider female participants, see Figure 3.4 for mean attractiveness ratings provided by female observers. A 2 (Sex: same, opposite) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA was conducted to investigate attractiveness ratings. A main effect of Sex was uncovered,

$F(1,31) = 15.5$, $p < .001$, $\eta^2 = .33$, whereby female participants considered people of the same sex to be more attractive than the opposite sex. A main effect of Familiarity was found, $F(1,31) = 29.9$, $p < .001$, $\eta^2 = .49$, indicating higher attractiveness ratings for familiar than unfamiliar people. An interaction was identified between these factors, $F(1,31) = 10.9$, $p = .002$, $\eta^2 = .26$, and Post-hoc Bonferroni adjusted *t*-tests with an alpha level of .013 (.05/4) were computed to investigate this interactive effect. Female participants considered familiar males to be more attractive than unfamiliar males, $t(34) = 5.61$, $p < .001$, $d = .95$. Furthermore, unfamiliar females were rated as more attractive than unfamiliar males, $t(34) = 4.01$, $p < .001$, $d = .68$. No other comparisons reached statistical significance, $p > .02$.

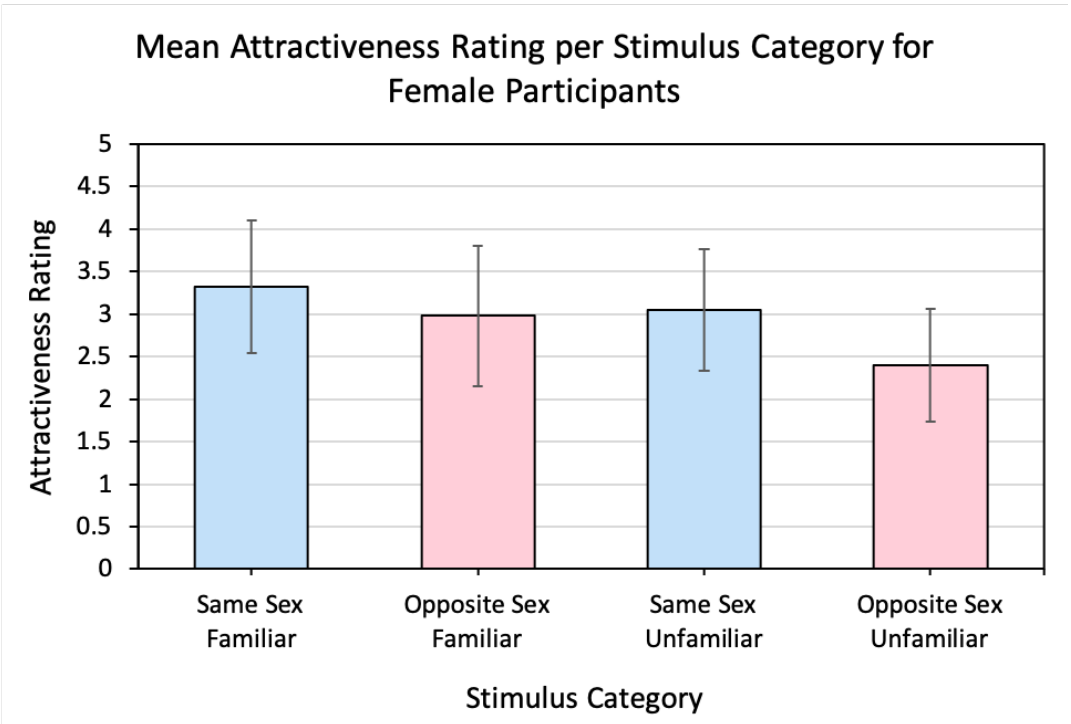

Figure 3.4. A bar graph representing mean attractiveness ratings provided by female participants per stimulus category with error bars indicating standard deviation.

Viewing Times

A 2 (Sex: same, opposite) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA was then computed for viewing times, see Figure 3.5 for average viewing length for female participants. A main effect of Sex was identified, $F(1,31) = 8.03$, $p = .008$, $\eta^2 = .21$, whereby females viewed people of the same sex for longer than the opposite sex. No main effect of familiarity was found, $F(1,31) = 0.07$, $p = .79$, $\eta^2 = .00$. Finally, there was no interaction present between Sex and Familiarity, $F(1,31) = 2.49$, $p = .12$, $\eta^2 = .07$.

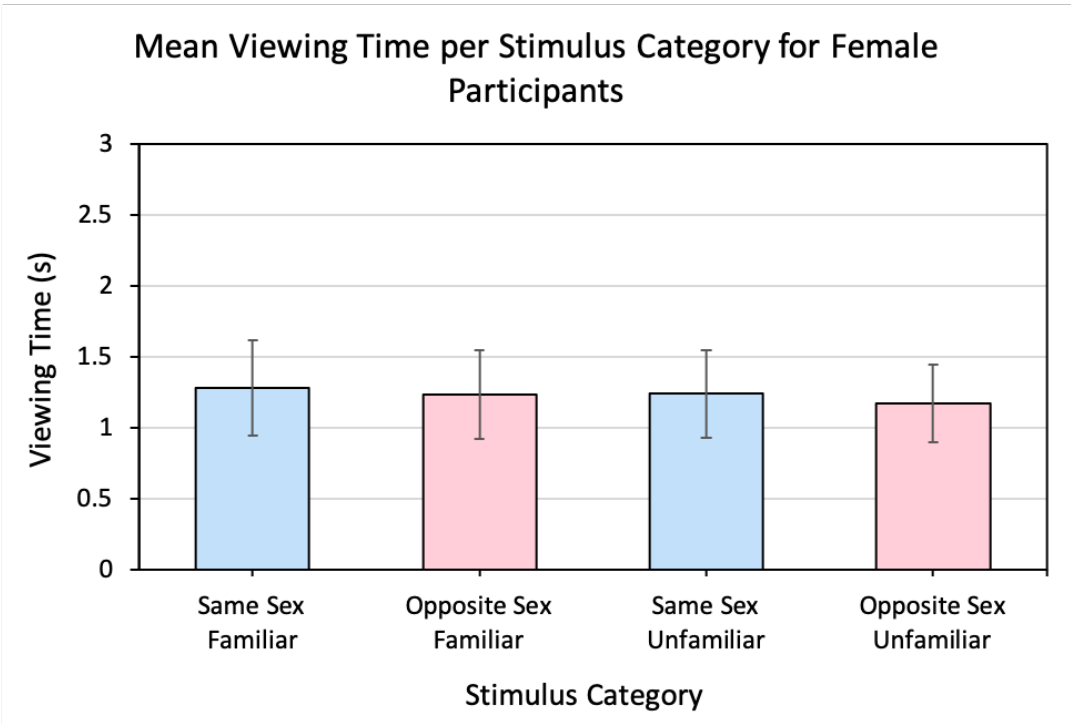

Figure 3.5. Bar graph featuring mean viewing times per stimulus categories for female observers. Standard deviation is presented through error bars.

Discussion

In this experiment, the aim was to re-examine the ability of cognitive measures, in this case viewing time, for measuring category-based attractiveness. When rating stimuli based on attractiveness, heterosexual male participants considered opposite sex images to be more attractive than same sex images. However, heterosexual female observers rated same sex stimuli as the most attractive. Additionally, female participants rated familiar models as more attractive than unfamiliar models, unlike male observers who exhibited similar ratings regardless of familiarity level. This suggests that explicit self-report tasks can measure attractiveness differences between sex groups within male observers. Female participants show a more complex response pattern, a result found throughout the wider literature across a range of methodologies (Bailey, 2009; Chivers et al., 2004).

When considering viewing time, results were not reflective of sexual orientation or familiarity with the stimulus model. The low correspondence between the self-report and the viewing time task suggests that attractiveness and familiarity may not be able to be accurately captured through implicit cognitive measurement. However, the absence of any viewing time effects could be due to the nature of facial stimuli, which may not be salient enough to produce an effective measure of attractiveness. This suggestion is supported by the lack of between sex category differences, the presence of which has been found consistently throughout the literature (Imhoff et al., 2010; Quinsey et al., 1996; Rönspies et al., 2015). To address this, subsequent experiments will employ more sexually salient stimuli, through use of full body imagery, to further investigate implicit measurement of attractiveness and familiarity.

In this experiment, familiarity with the stimulus model was investigated to explore any potential interaction with perceived attractiveness, and to further inform the possible applications of viewing time as a measure. Despite research suggesting familiarity can lead to higher ratings of attractiveness (Han et al., 2020; Moreland &

Beach, 1992; Rhodes et al., 2005), this was only identified within this experiment when considering female observers evaluations of male targets.

Experiment 3

Experiment 2 provided a basis of understanding of how attractiveness and familiarity may interact between sex groups and what stimuli may inhibit the cognitive measurement of these factors using cropped faces as stimuli. Experiment 3 was designed to replicate these findings while employing a full body stimulus set, thus addressing the caveat introduced by the facial stimuli in Experiment 2, whereby facial images did not elicit viewing time differences reflective of sexual orientation. Furthermore, this stimulus set allowed for the exploration of attractiveness effects both *between* and *within* sex categories. Images of familiar and unfamiliar male and female celebrities were presented to participants, with within-person variability introduced to limit the influence of identity confounds. To this end, each person was represented by two images, one wearing casual clothing and one in beachwear, in order to vary attractiveness across each identity. Therefore, Experiment 3 was designed to build upon the findings of Experiment 2 and to provide the first evidence of empirical investigation aiming to systematically examine both between and within sex category differences in attractiveness alongside familiarity.

Method

Participants

Eighty-one participants, 40 male (M_{age} = 44.76 years, *SD* = 15.56, Range = 21 to 77) and 41 females, (M_{age} = 40.83 years, *SD* = 12.98, Range = 20 to 66), took part in this experiment online through Prolific. All participants self-identified as heterosexual using the Kinsey scale (Kinsey et al., 1948), and resided in the UK at time of participation.

Materials

The stimuli consisted of a total of 80 images, comprising of 40 female and 40 male celebrities. This aligns with the stimulus numbers often employed throughout the literature (Imhoff et al., 2010; Lippa, 2012; Lippa et al., 2010). Within each gender category, there were 20 identities – 10 familiar celebrities from the UK or US, and 10 from Australia. Two images were included for each celebrity depicting this person in casual wear and swimwear. The images were sourced from Google images, and resized to 600 pixels in height, with no constraint on width. The backgrounds were edited out and replaced with a grey screen. Example stimuli are provided in Figure 3.6.

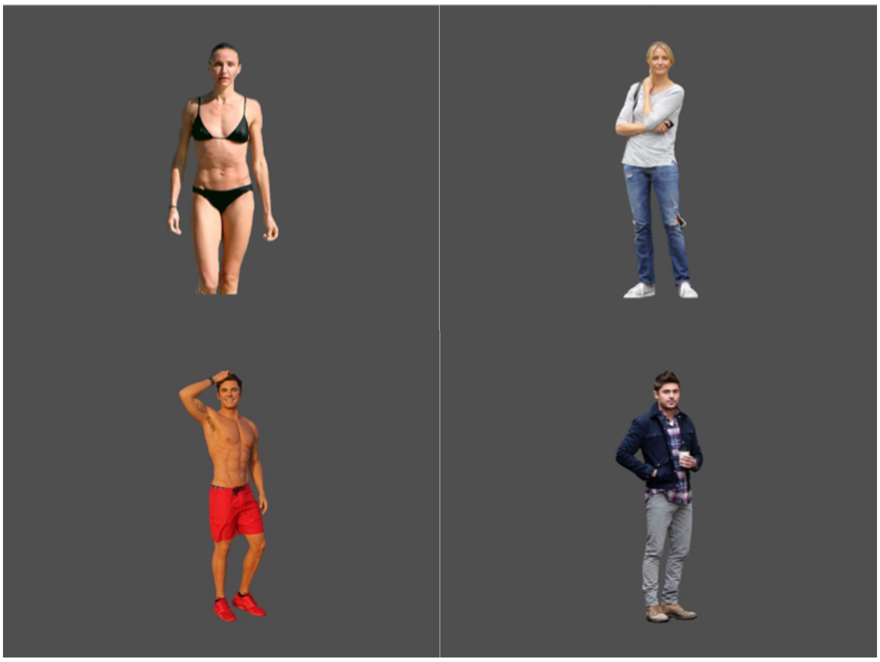

Figure 3.6. Example stimuli for this experiment, depicting a female and male celebrity in swimwear (left) and casual wear (right).

Procedure

This experiment was programmed and implemented online in the same way as Experiment 2. Participants were informed that they would be viewing images of celebrities in casual clothing and swimwear and were instructed to rate stimuli based on how attractive they find them. Ratings were given according to a 5-point Likert scale, with the scores ranging between 1 representing '*not attractive'* and 5 *'highly attractive'*. This attractiveness task was separated into four identical blocks, with all images shown once in each. Participants were offered a break between blocks two and three. The time taken to provide an attractiveness response was collected for each stimulus. After the completion of this task, a final block was provided, constituting a familiarity check. All stimuli were again presented, this time with instructions to indicate how familiar the participant was with the image model. Ratings were again taken on a five-point scale, ranging from 1 *'unfamiliar'* to 5 '*highly familiar'*.

Results

To ensure that the familiar and unfamiliar stimulus categories were capturing participant familiarity rather than assumed familiarity, the data was filtered in the same process as Experiment 2, thus allowing for accurate exploration of this factor. This resulted in the exclusion of 46.30% of trials featuring familiar celebrities and 25.80% of those presenting unfamiliar celebrities. Post-hoc power analysis indicated a statistical power of .20 for this experiment. Akin to the previous experiment, an outlier criteria of 6 seconds was included to minimise outlier effects.

Male Participants

Ratings

Firstly, the attractiveness ratings provided by male participants were considered. Figure 3.7 represents attractiveness ratings for male observers and shows higher judgements for opposite sex than same sex stimuli across all clothing and familiarity conditions.

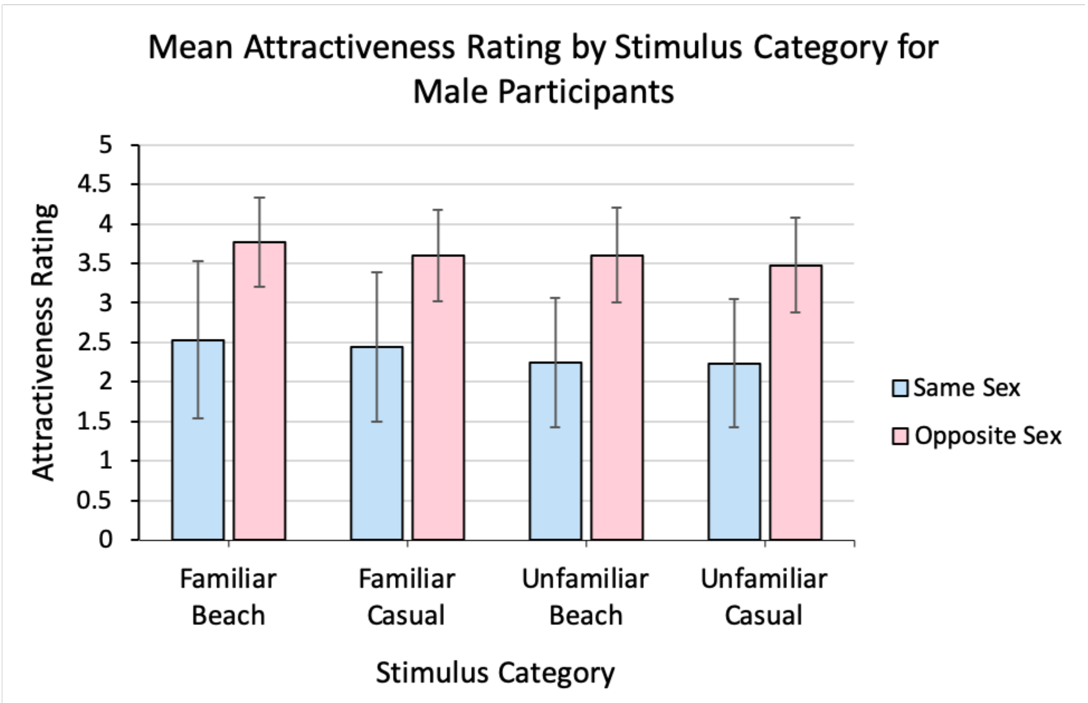

Figure 3.7. Bar graph showing attractiveness ratings per stimulus category for male participants. Error bars indicate standard deviation.

A 2 (Sex: same, opposite) x 2 (Attractiveness: beach, casual) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA showed a main effect of Sex, $F(1,32) = 74.03$, $p < .001$, $\eta^2 = .42$, whereby male participants rated images featuring women as more attractive than images of men. A main effect of Familiarity was also found, $F(1,32) = 14.46$, $p < .001$, $\eta^2 = .01$, due to familiar others being rated as more attractive. Furthermore, a main effect of Attractiveness, $F(1,32) = 6.87$, $p = .01$, $\eta^2 =$

.004, was uncovered due to higher attractiveness ratings for people in beach than casual wear. There were no interactions present between Sex and Attractiveness, $F(1,32) = 1.38$, $p = .25$, $\eta^2 = .001$, Sex and Familiarity, $F(1,32) = 1.04$, $p = .32$, $\eta^2 = .001$, or Attractiveness and Familiarity, $F(1,32) = 0.71$, $p = .42$, $\eta^2 = .00$, and no 3-way interaction between these factors, $F(1,32) = 0.63$, $p = 43$, $\eta^2 = .00$.

Viewing Time

A 2 (Sex: same, opposite) x 2 (Attractiveness: beach, casual) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA was conducted for viewing time within male participants, which is illustrated in Figure 3.8.

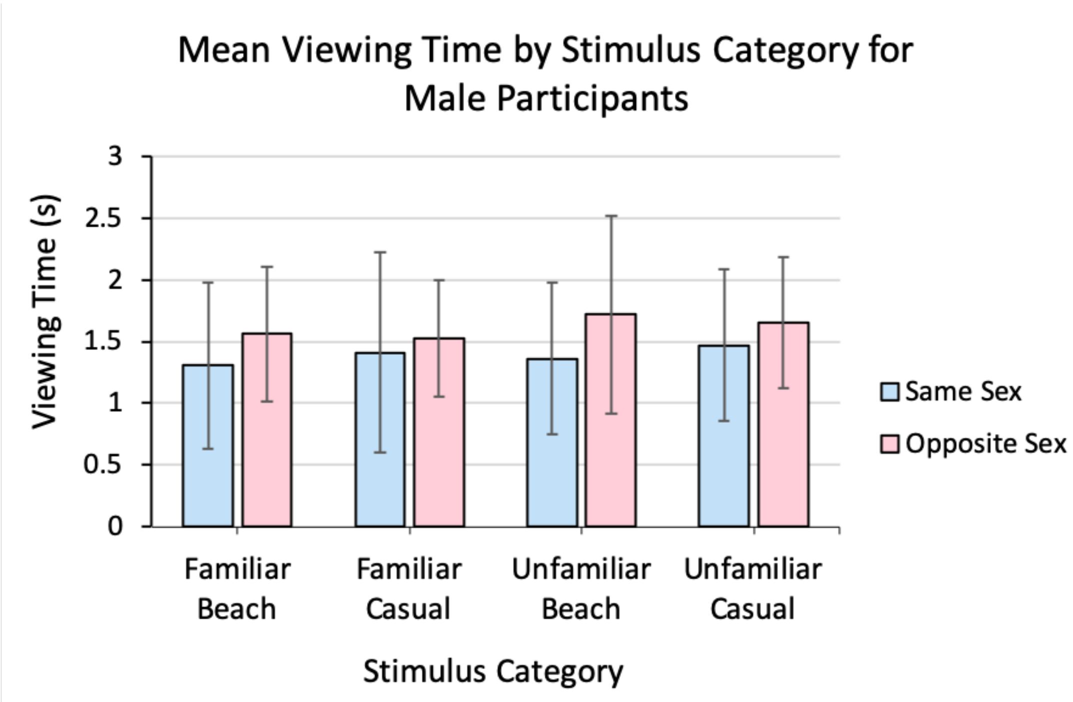

Figure 3.8. Bar graph featuring viewing time per stimulus category for male participants. Error bars indicate standard deviation.

This ANOVA revealed a main effect of Sex, $F(1,22) = 8.56$, $p = .008$, $\eta^2 = .04$, due to longer viewing times for people of the opposite sex to the observer than the same sex. There were no main effects of Attractiveness, $F(1,22) = 0.03$, $p = .87$, $\eta^2 = .00$, or Familiarity, $F(1,22) = 0.002$, $p = .97$, $\eta^2 = .00$, and no interactions between Sex and Attractiveness, $F(1,22) = 3.78$, $p = .07$, $\eta^2 = .004$, Sex and Familiarity, $F(1,22) = 0.55$, $p = .47$, $\eta^2 = .002$, Attractiveness and Familiarity, $F(1,22) = 0.01$, $p = .91$, $\eta^2 = .00$, or all three factors, $F(1,22) = 0.02$, $p = .90$, $\eta^2 = .00$.

Female Participants

Ratings

Finally, the data was split to consider only female participants. The attractiveness ratings were considered first, with mean ratings for all conditions represented in Figure 3.9.

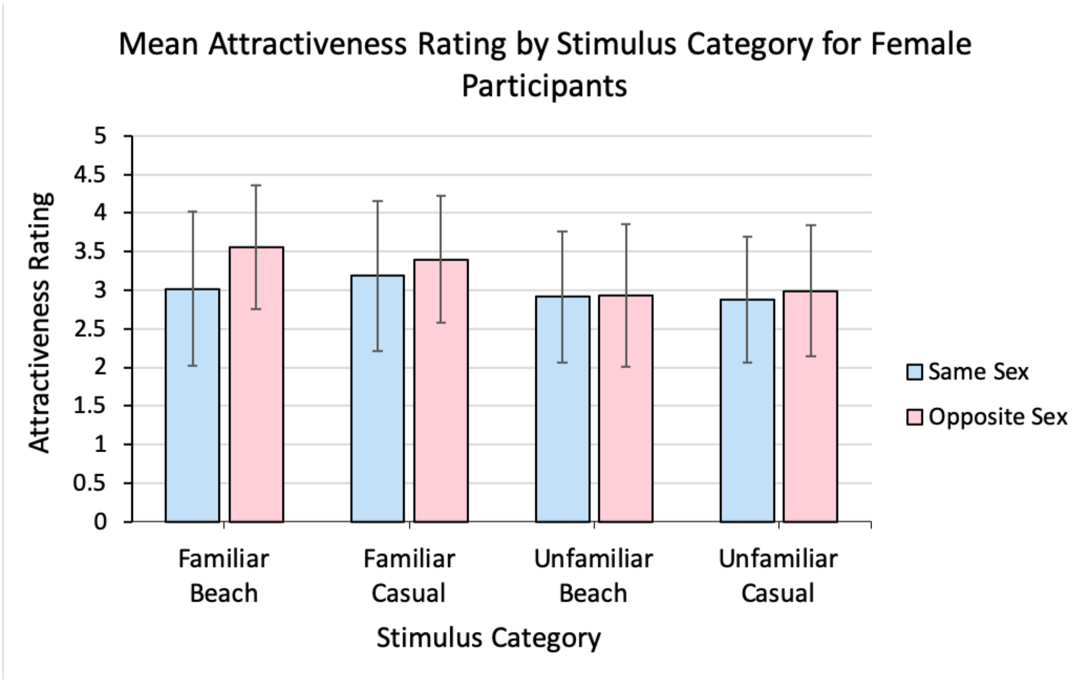

Figure 3.9. Bar graph showing mean attractiveness ratings per stimulus category for female participants with error bars indicating standard deviation.

A repeated-measures 2 (Sex: same, opposite) x 2 (Attractiveness: beach, casual) x 2 (Familiarity: familiar, unfamiliar) ANOVA indicated a 3-way interaction of Sex, Attractiveness and Familiarity, $F(1,33) = 11.79$, $p = .002$, $\eta^2 = .002$. Consequently, data was split by familiarity to further analyse this interaction and two additional ANOVAs conducted.

Familiar people were considered first, with a 2 (Sex: same, opposite) x 2 (Attractiveness: beach, casual) ANOVA showing no main effect of Attractiveness, $F(1,37) = 0.009$, $p = .92$, $\eta^2 = .00$, but a main effect of Sex, $F(1,37) = 17.86$, $p < .001$, $\eta^2 = .04$, due to higher attractiveness ratings for opposite sex than same sex images. Additionally, an interaction between Sex and Attractiveness was identified, $F(1,37) = 8.55$, $p = .006$, $\eta^2 = .01$. Bonferroni adjusted *t*-tests, with an alpha level of .013 (.05/4), were conducted to investigate this effect and indicated that people in beachwear were rated as more attractive when these were of the opposite sex of the observer than of the same sex, $t(38) = 5.61$, $p < .001$, $d = .90$. Beachwear images were considered marginally more attractive than casualwear stimuli within people of the opposite sex as the participant, $t(39) = 2.23$, $p = .03$, $d = .35$. Furthermore, casual images of the opposite sex were considered marginally more attractive than same sex stimuli, $t(38) = 2.09$, $p = .04$, $d = .34$. Finally, people in casualwear were rated as similarly attractive as those in beachwear when of the same sex as the observer, $t(37) = 1.44$, $p = .16$, $d = .23$. $t(38) = 2.09$, $p = .04$, $d = .34$.

An equivalent 2 (Sex: same, opposite) x 2 (Attractiveness: beach, casual) repeated-measures ANOVA was computed for unfamiliar stimuli. This did not show a main effect of Sex, $F(1,36) = 0.91$, $p = .35$, $\eta^2 = .001$, or of Attractiveness, $F(1,36) = 0.04$, $p = .83$, $\eta^2 = .00$, and no interaction between these factors, $F(1,36) = 1.08$, p

= .31, $\eta^2 = .001$. This indicates that the 3-way interaction found between Sex, Attractiveness and Familiarity is driven by the stimuli depicting familiar people.

Viewing Time

A final 2 (Sex: same, opposite) x 2 (Attractiveness: beach, casual) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA was conducted to investigate viewing time within female observers. The means for this participant group are illustrated in Figure 3.10 and show comparable viewing times across all conditions.

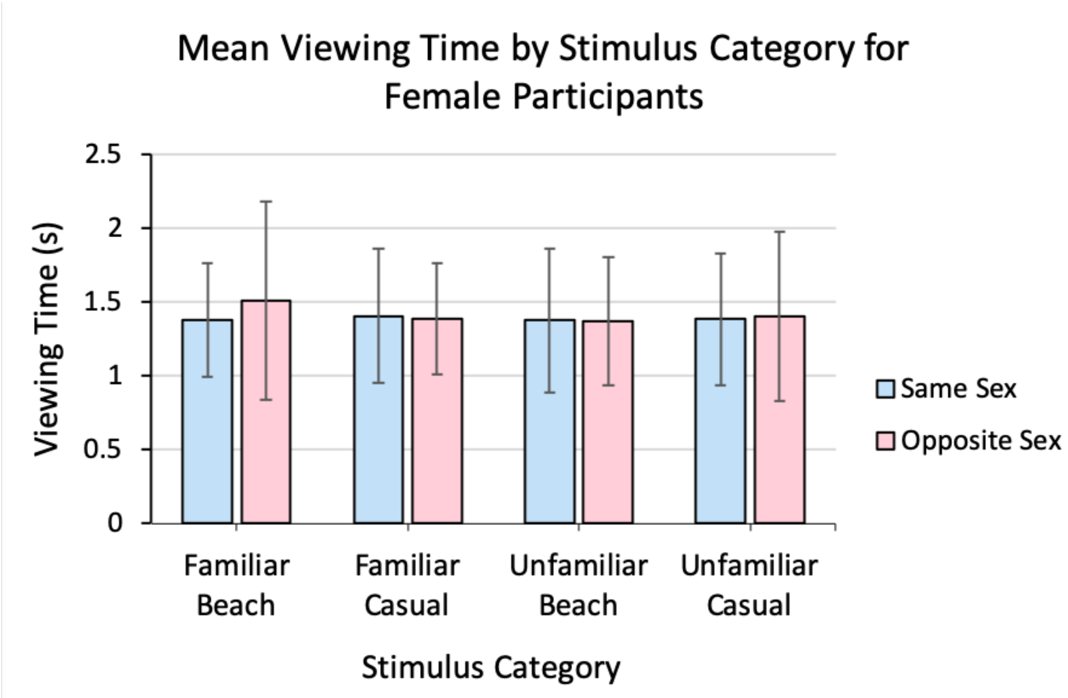

Figure 3.10. Bar graph representing viewing time per stimulus category for female participants, the error bars indicate standard deviation.

Accordingly, no main effects of Sex, $F(1,25) = 0.73$, $p = .40$, $\eta^2 = .002$, Attractiveness, $F(1,25) = 1.01$, $p = .33$, $\eta^2 = .002$, or Familiarity, $F(1,25) = 2.44$, $p = .13$, $\eta^2 = .003$, were discovered. Similarly, there were no interactions between Sex and Attractiveness, $F(1,22) = 0.56$, $p = .46$, $\eta^2 = .001$, Sex and Familiarity, $F(1,22) = 2.34$, $p = .14$, $\eta^2 = .003$, Attractiveness and Familiarity, $F(1,22) = 2.27$, $p = .14$, $\eta^2 = .004$, and between all three factors, $F(1,22) = 0.11$ $p = .75$, $\eta^2 = .00$.

Discussion

The aim of this experiment was to explore the sensitivity of implicit cognitive tasks for the measurement of *between* and *within* sex category attractiveness. Male observers clearly elicited a sex category-based effect, with opposite sex images leading to longer viewing times than same sex stimuli. However, female participants did not show distinguishable viewing time differences for stimulus sex, mirroring the complex nature of female sexuality identified throughout the literature (Bailey, 2009; Chivers et al., 2004, 2007).

More importantly, the sensitivity of viewing time was considered in relation to within-category responses. As such, participant data for each stimuli context was analysed, rather than the broader categories of male and female. This attractiveness concept was explored within this experiment by providing a more revealing image (swimwear) and a less revealing image (casual clothing) of each identity. Beach images were rated as more attractive than casual images by the male participants while female observers did not indicate a significant difference in attractiveness across these contexts. However, it should be noted that across all conditions, ratings fell between two and four on a five-point scale. Potential effects may therefore have been masked by these relatively similarly attractive stimuli. Due to the absence of ceiling effects, it can be assumed that these ratings were sensitive to real response and not measurement error, and thus these stimuli were genuinely considered similarly attractive. Despite significant differences identified via rating scores, there were no differences in viewing times across these images. This would suggest that viewing time may not be reliably sensitive when considering attractiveness *within* sex categories. Furthermore, no influence of familiarity was uncovered within viewing time. This finding emulates Experiment 3 and further suggests that implicit cognitive measures are not able to accurately investigate familiarity.

Experiment 4

Experiment 3 indicated that differences in attractiveness within sex categories can be measured through explicit self-report measures. Additionally, when employing an implicit measure of viewing time, longer viewing times are present when observing opposite sex stimuli. However, the more fine-grained differences in attractiveness within sex category, as measured through within-person variation, were not indicated in an implicit viewing time task. These responses are specific for male participants, with Experiments 2 and 3 showing that females exhibit more complex and inconsistent responses, as is often identified when exploring female sexuality (Bailey, 2009; Chivers et al., 2004; Quinsey et al., 1996). As participant sex was identified to affect results, as can be expected from previous studies, a follow-up experiment was then conducted to isolate effects of attractiveness and familiarity when this factor is eliminated from the design. This also serves to ensure that the strong categorical decision of sex is not obscuring the more subtle differences present in familiarity and within-sex attractiveness. By eliminating this factor from the design, it is possible that these effects will be enhanced. If viewing time measures exclusively reflect attractiveness to sex category, then there should be no difference in viewing time length between casually clothed and beachwear stimuli.

However, if more fine-grained attractiveness distinctions are encapsulated by this response, then longer viewing times should be present when viewing the more attractive swimwear stimuli.

Method

Participant, Materials and Procedure

This experiment was conducted online through Prolific. Sixty-one heterosexual male participants, all of whom identified as heterosexual using the Kinsey scale (Kinsey et al., 1948), who held residence in the UK initially took part in this experiment (M_{age} = 35.52, *SD* = 14.95, Range = 19 to 73). Fourteen of these aborted the experiment before completion, leading to analysis of 47 data files. The stimuli and procedure were similar to Experiment 3 with the exclusion of female participants and male stimuli.

Results

All analyses were conducted on data which was filtered based on familiarity, as outlined in Experiment 2. This resulted in the exclusion of 46.60% of trials featuring familiar celebrities and 41.06% of those presenting unfamiliar celebrities. Post-hoc statistical power was calculated at .30. Furthermore, a 6 second exclusion criteria was calculated to limit outlier effects, consistent with the previous experiments.

Ratings

In a first step of the analysis, attractiveness ratings to images of women in beach and casualwear were analysed. The mean values for all stimuli categories are shown in Figure 3.11.

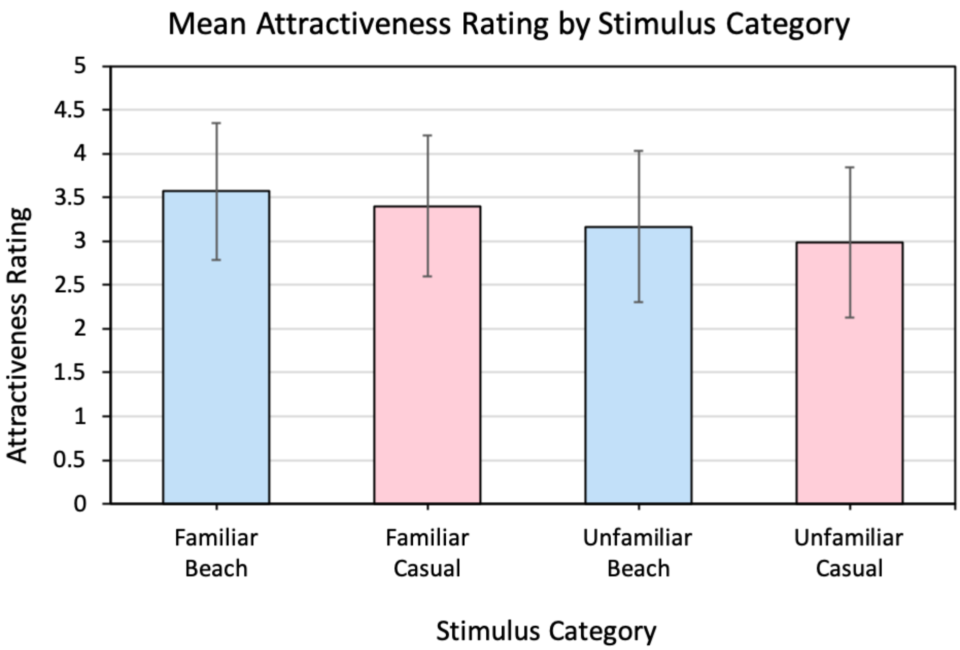

Figure 3.11. Bar graph illustrating mean attractiveness rating for each category of stimuli with error bars representing standard deviation.

A 2 (Attractiveness: beach, casual) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA of these data revealed a main effect of Attractiveness, $F(1,33) = 5.56$, $p = .02$, $\eta^2 = .01$, reflecting higher attractiveness ratings for people in beach than casualwear. A main effect of Familiarity was also found, $F(1,33) = 6.40$, $p = .02$, $\eta^2 = .04$, indicating that familiar people were rated as more attractive than unfamiliar people. The interaction of Attractiveness and Familiarity was not significant, $F(1,33) = 0.02$, $p = .90$, $\eta^2 = .00$.

Viewing Time

A 2 (Attractiveness: beach, casual) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA was then conducted to analyse viewing time. The means for this data are presented in Figure 3.12. For this data, no main effects of Attractiveness, $F(1,33) = 2.11$, $p = .16$, $\eta^2 = .02$, or Familiarity were found, $F(1,33) =$

0.08, p = .78, η^2 = .00, and no interaction between these factors, $F(1,33)$ = 0.61, p = .44, η^2 =.002.

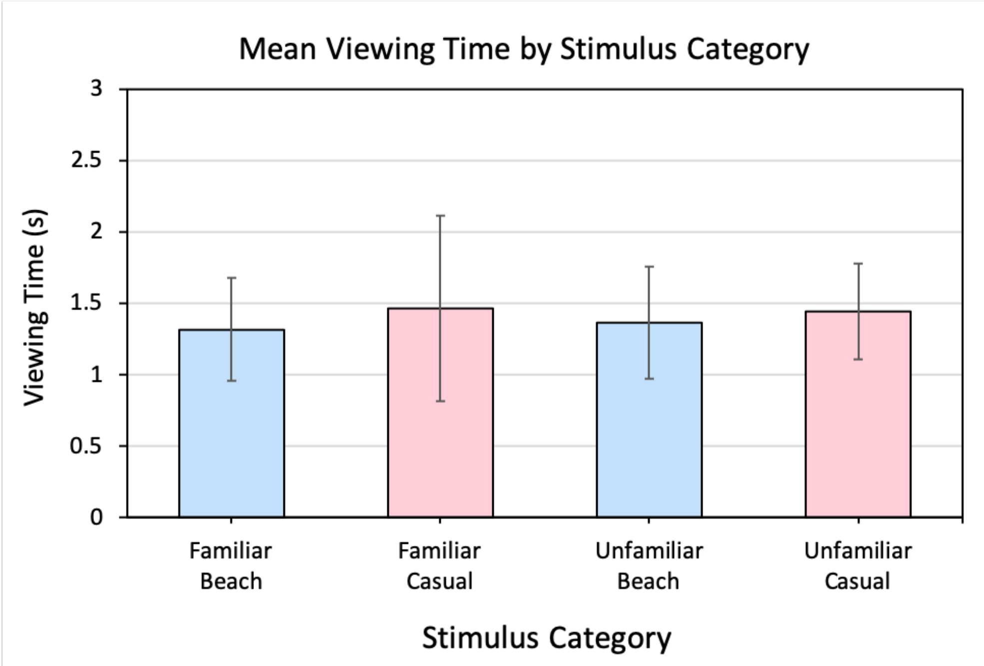

Figure 3.12. Mean viewing time length for each stimulus category. Error bars illustrate standard deviation.

Discussion

Experiment 3 examined attractiveness both between and within sex category and revealed that viewing time effects were only present in male observers. This experiment therefore aimed to further investigate attractiveness at the within-category level. To this end, the experimental design introduced in Experiment 3 was replicated, with the absence of female participants and male stimuli. The results corroborated those from Experiment 3, with higher explicit attractiveness ratings recorded for people in beach than casual wear. Furthermore, familiar people were considered more attractive than unfamiliar others. However, when considering the implicit viewing time measure, no differences were identified between stimuli context (beach vs. casual), or familiarity. As such, this experiment supports the suggestion that this cognitive measure is not sensitive to attractiveness differences *within* sex categories.

Experiment 5

The preceding experiments demonstrate consistently that viewing time may not be sensitive to within-category differences between images of the same persons, such as people depicted in beach and casualwear. In Experiment 5, the generalisability of these effects is examined with a second cognitive measure that has been applied extensively in the study of cognition, namely the Pictorial Stroop task (Ashwin et al., 2006; Constantine et al., 2001; Gallagher-Duffy et al., 2009). When utilising this task for the investigation of sexual attraction, images of people are presented with an overlay of colour (Ciardha & Gormley, 2009; Mannfolk et al., 2023), with experiments showing both that response time reflects sexual orientation, and age-related interest (Ciardha & Gormley, 2012). As such, the Pictorial Stroop will be utilised to assess participants' impressions of stimulus model attractiveness in the current experiment. To this end, the same stimuli used in experiments two and three were altered to provide the aforementioned colour tint. Akin to prior experiments, familiarity was also investigated. This experiment therefore aimed to explore the sensitivity of the Pictorial Stroop in investigating attractiveness of and familiarity towards within sex category stimuli.

Method

Participants

This experiment was conducted online through Prolific and included 56 male participants, with a mean age of 37.93 (*SD* = 15.80, Range = 18 to 75). All participants self-identified as heterosexual according to the Kinsey scale (Kinsey et al., 1948), and resided in the UK at time of testing. Six participants were excluded from data analysis due to experiment termination before completion.

Materials

This experiment employed the same stimuli as Experiment 3, and thus included 40 images featuring female celebrities. Half of these celebrities were from the UK/ America and the other half from Australia. Two images per identity were included, one in beachwear and another in casual clothing. Each image measured 600 pixels in height, with width left unconstrained to avoid distortion effects. Image background was cropped and replaced with a solid grey screen.

To facilitate the pictorial Stroop task, each image was modified to include a transparent overlay of yellow, green, blue and red, see Figure 3.13. Four conditions were employed, with each colour image being presented once, so that no two conditions had the same image and colour combination. The ratio of each colour was consistent across conditions. Each of the images within every condition was shown four times, with an initial practice block of plain colours. Images remained on screen until a colour indication was made.

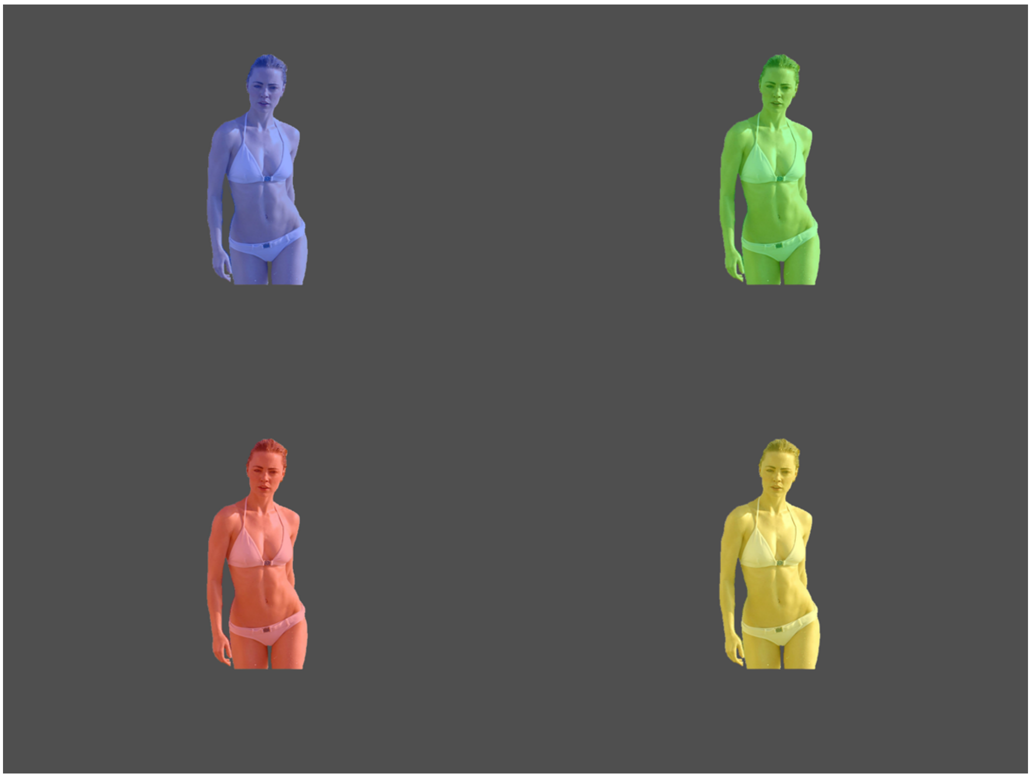

Figure 3.13. An array of example stimuli featuring the same unfamiliar identity, displaying the four different colour overlays. The top row shows the beach stimuli, with the bottom row showing the casual set.

Procedure

Experiment 5 was distributed in the same way as Experiments 2, 3 and 4. Participants were first presented with a practice block containing 80 trials of blank-coloured rectangles. This was provided to ensure participants ample time to practice the response keys. Following this block, four blocks containing the stimuli were presented. Each block consisted of the same images and were representative of one of the aforementioned conditions. The response keys were as follows: 1 for blue, 2 for green, 3 for red and 4 for yellow. One final block was presented consisting of all images without a colour hue overlay. Participants were asked to indicate on a 5-point scale their familiarity with the stimulus model. This served to help aid data analysis and to tie this experiment in with the wider study project.

Results

All data was filtered to account for participant familiarity with stimuli identity, in keeping with the previous experiments. Only stimuli which participants were familiar with were analysed from the familiar stimulus category, and only unfamiliar identities were included from the unfamiliar category. This facilitates accurate investigation into familiarity and its potential interactions with attractiveness within cognitive measures. This resulted in the exclusion of 50.50% of trials featuring familiar celebrities and 45.70% of those presenting unfamiliar celebrities. The post-hoc statistical power for this experiment equalled .46. In line with the previous experiments, a 6 second outlier criteria was applied for this Stroop task.

Accuracy

Firstly, accuracy was explored for each of the stimulus categories to see if the content of the stimuli impacted the accuracy of colour identification. The mean percentage accuracy scores for these conditions are illustrated in Figure 3.14. A 2 (Attractiveness: beach, casual) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA of this data was conducted. This did not show a main effect of Familiarity, $F(1,49) = 0.07$, $p = .80$, $\eta^2 = .00$, but revealed a marginal main effect of Attractiveness, $F(1,49) = 4.52$, $p = .04$, $\eta^2 = .003$, reflecting higher accuracy for beach images than casualwear stimuli. In addition, an interaction between Attractiveness and Familiarity was identified, $F(1,49) = 7.63$, $p = .008$, $\eta^2 = .004$. To investigate this effect, *t*-tests with a Bonferroni adjusted alpha level of .013 (.05/4) were conducted, which indicated that this effect was driven by the differences between the familiar beach and familiar casual stimulus categories, $t(49) = 3.15$, $p = .003$, $d = .45$, with the other comparisons of unfamiliar beach and unfamiliar casual, $t(49) = 0.21$, $p = .83$, $d = .03$, familiar beach and unfamiliar beach, $t(49) = 1.93$, $p = .06$, $d = .27$, and familiar casual and unfamiliar casual, $t(49) = 1.45$, $p = .15$, $d = .21$, not reaching statistical significance.

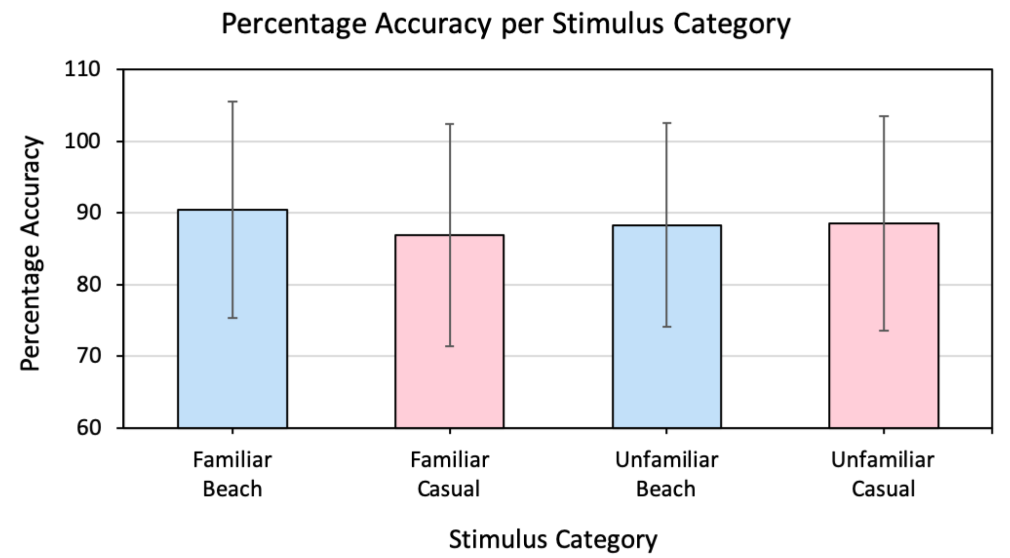

Figure 3.14. Bar graph showing mean accuracy percentage per stimulus category with error bars representative of standard deviation.

Response Time

A 2 (Attractiveness: beach, casual) x 2 (Familiarity: familiar, unfamiliar) repeated-measures ANOVA was then conducted to analyse time taken to make an accurate colour response. These data are illustrated in Figure 3.15. No main effect of Attractiveness, $F(1,49) = 0.18$, $p = .67$, $\eta^2 = .00$, or Familiarity was found, $F(1,49) = 2.48$, $p = .12$, $\eta^2 = .01$, and there was no interaction between these factors, $F(1,49) = 1.70$, $p = .20$, $\eta^2 = .003$.

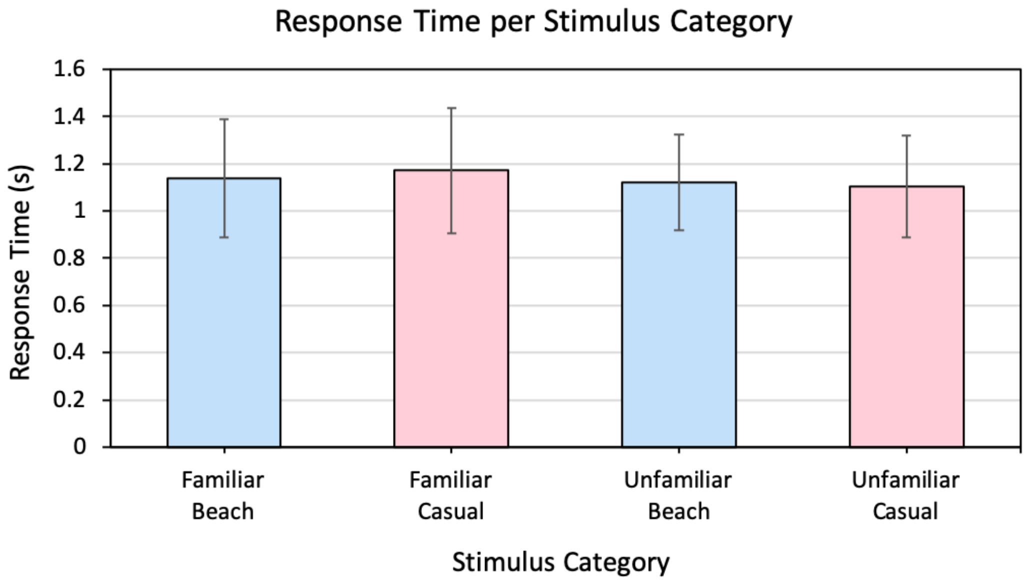

Figure 3.15. A bar graph depicting mean response times per stimulus category, with error bars indicating standard deviation.

Discussion

Overall accuracy in indicating colour tints for each stimuli category fell between 85-90%. High accuracy suggests that participants were adhering correctly to instructions and that task difficulty was manageable. Response accuracy between each stimulus type was comparable, with the only significant difference occurring between the familiar beach and familiar causal images, with higher accuracy for the beach images. However, with all stimulus types showing high and relatively similar scores across the four categories, it is unlikely that accuracy exerted influence over reaction times in this experiment. Accordingly, no main effect of attractiveness was discovered in reaction times, mirroring the previous viewing time experiments. This indicates that, akin to viewing time, Pictorial Stroop tasks may not be sensitive to within-category differences in attractiveness. Furthermore, attractiveness and familiarity did not interact in this experiment, suggesting that these factors do not exert influence over one another during cognitive measurement. When considering familiarity, comparable responses were observed for familiar and unfamiliar others. In general, these results highlight those found throughout this chapter, whereby within sex category attractiveness and familiarity did not exert clear and consistent effects within implicit measurement, unlike the wider category of between-sex attraction.

General Discussion

Across the experiments in this chapter, cognitive latency-based measures were reflective of observers sexual orientation, whereby response times were longer to images of the opposite sex than the same sex as the observer. However, these effects were specific to male participants whilst viewing full body images (Experiments 3-5) and were not identified for female observers (Experiment 2-3) or images in which only the face is shown (Experiment 2). Moreover, when exploring attraction within sex categories, it was found that implicit measures were not sensitive to finer grained differences in attraction, such as different images of the same person that vary in their perceived attractiveness (Experiments 3-5). This was found consistently across various viewing time paradigms, a pictorial Stroop task and face and full body stimulus sets.

Experiment 2 did not show the typical between sex category differences expected when utilising cognitive measures, instead no sex category differences were found for facial images. This offers insight into what is being measured through these tasks. Experiment 3 found these sex category attractiveness differences when employing full body images, stimuli which are more sexually salient than faces alone. This suggests that cognitive measures of attractiveness may be tapping specifically into mate preference rather than simple attractiveness. This assumption is supported further when considering that cognitive tasks can reflect both sexual orientation (Harris et al., 1996; Quinsey et al., 1996; Rieger & Savin-Williams, 2012), and age-related interest (Attard-Johnson et al., 2016; Ciardha & Gormley, 2012), two concepts which act to direct attention towards potential mate groups. Therefore, it is plausible that cognitive measures are capable of capturing mate preference, but are not able to distinguish between individuals belonging to the same group, such as within sex categories.

When considering the full body stimuli utilised throughout the rest of the study, results showing that longer viewing times for opposite sex images within male participants converge with other studies that have identified similar effects. Throughout the literature, longer viewing times have been discovered for heterosexual male participants when viewing images of women, with female participants diverging from this pattern and showing less category-specific responses (Quinsey et al., 1996; Wright & Adams, 1999). While male sexual response appears to orient towards mating opportunities, female responses are more fluid, with heterosexual women often showing bisexual response patterns (Bailey, 2009). Indeed, women show complicated sexual responses across a range of biological and cognitive measures (Attard-Johnson et al., 2021; Chivers et al., 2004; Peterson et al., 2010).

Whilst these between sex category attractiveness effects have seen frequent investigation, within-category differences have not seen such extensive review. Despite this, there have been a few experiments which have considered these high-level attractiveness effects alongside the main category findings (Israel & Strassberg, 2009; Lippa et al., 2010). However, there have been no conclusive findings across this literature. There has been some suggestion that within-category attractiveness can be identified using implicit cognitive measures. Specifically, within male participants, Lippa (2012) found that increased ratings of attractiveness lead to longer viewing times when viewing female models. Women showed the expected bisexual viewing time pattern. However, regardless of stimulus sex, female participants exhibited longer viewing times the more attractive the image.

In contrast, the current study did not identify any within-category effects of attraction. Synonymous with this finding, some evidence suggests the limits of cognitive measures in distinguishing fine-grained levels of attractiveness. One such experiment (Israel & Strassberg, 2009) suggested that the short length of time for which images are presented in viewing time experiments does not allow for sensitivity to the moderate attractiveness differences between stimuli within sex categories. Prior to the current study, such within-category distinctions were often adjacent to the main hypotheses of the experiments. This study is the first of its kind to prioritise such investigation and demonstrates that despite clear between-category effects, cognitive measures of attractiveness are not sensitive to fine-grained levels of attractiveness.

The results presented in this study suggest that cognitive measures of attractiveness are sensitive only to differences *between* sex-categories, and not to the finer grained attractiveness present *within* such categories. Thus, results suggest either that these paradigms do not measure beauty or that beauty may not be skin deep, and rather can transcend the superficial differences between potential mates when measuring at an implicit level. It is clear that, at a conscious level, attractiveness is perceived to differ across identity (Jenkins et al., 2011). However, results from this study suggest that attractiveness measured implicitly may be driven by general mate preference and not the subtle variations between or within individuals of the preferred sex. However, it should be noted that the statistical power of the experiments in this chapter were relatively low. As such, caution should be taken when drawing conclusions based on the limited sample sizes included. These studies would benefit from direct replication with a largely increased participant population, as this would increase confidence in any identified findings.

Throughout this study, within sex category attractiveness was investigated whilst controlling for identity. When considering within-category attractiveness, differences in appearance exist not only between people from the same gender category, but within each individual (Burton, 2013; Jenkins et al., 2011). As such, the attractiveness of an individual varies across different instances of identity, thus calling into question the potential role of familiarity on attractiveness. Previous research utilising cognitive measures have observed that as familiarity with an individual increases, as does their perceived attractiveness (Han et al., 2020; Moreland & Beach, 1992; Rhodes et al., 2005). Therefore, it stands to reason that if participants are familiar, or perceive familiarity with, a stimulus model, this may impact attractiveness responses. In order to control for and investigate familiarity, identity was controlled in this study. In Experiments 2 to 4, two instances of each stimulus identity were presented – one higher (beachwear) and one lower (casual clothing) in explicit content. It was presumed that the more sexually explicit an image, the more attractive it is considered to be, an assumption corroborated by participant self-report ratings. Familiarity was measured in a quasi-design, with half of the stimuli models being famous individuals, i.e. celebrities from the United Kingdom and United States of America, with the other half famous only in Australia. This ensured that half of the stimuli set would likely be familiar to the participants, all of which were located in the UK. However, across the four included experiments, there were no consistent familiarity findings identified.

In summary, this study has systematically investigated the capabilities of implicit cognitive measures of attraction. Results suggest that, within male participants, implicit measures are sensitive to attraction *between* sex categories, but not reflective of differences in attractiveness *within* these categories. Moreover, such measures are not reflective of higher-level distinctions, such as familiarity. Therefore, it is plausible that implicit measures capture mate preference, which orientates attention towards a group of attractive others, rather than fine grained distinctions of familiarity or the subtle attractiveness differences present between and within individuals of the same sex category.

When considering the findings from this experiment, it must be taken into account that there remains a measure of disparity in the language used to describe this phenomenon. Across rating tasks in this research field, the actual indications participants are instructed to perform are often ratings of sexual attractiveness (Harris et al., 1996; Quinsey et al., 1996; Worling, 2006). However, this terminology is not exclusive, with some studies utilising other rating scales such as (general) attractiveness (Hassebrauck, 1998) or sexual appeal (Dawson et al., 2012; Ebsworth & Lalumière, 2012; Lalumière et al., 2018). Regardless, during analysis all results are amalgamated into male and female categories and the differences between groups compared. While all such experiments conclude measurement of the same underlying factor, the lack of standardisation in terminology undermines this assumption. Further experimental research to this effect will be required to address this potential caveat as without empirical research into what exactly such tasks can indicate, the full theoretical implications surrounding cognitive measures of attractiveness remain unexplored.

Furthermore, another factor that must be considered when comprehending the results of this study is the relatively high proportion of excluded trials. Across both familiar (M_{age} = 47.78%) and unfamiliar (M_{age} = 38.44%) stimuli, large amounts of trials were excluded as the included stimuli were indicated to belong to the opposite familiarity category. This calls into question the validity of these experiments as it is possible that a high degree of randomness was present when participants made such indications, due to these rates being close to chance. However, the stimulus models chosen for these experiments were all public figures. While it was believed that participants would be more familiar with celebrities predominantly involved with projects in their origin country (for example, British tv soap operas or films), many of these projects are available worldwide. Therefore, it is not improbable that participants are genuinely familiar with celebrities initially categorised as unfamiliar. As such, these trials were excluded from analysis to negate any effects driven by this possibility. Furthermore, it is possible that participants were unfamiliar with presumed familiar celebrities if they have never engaged with media containing these individuals. While repeating this study with everyday stimulus models that participants are personally familiar or unfamiliar with may an effective way to combat the potential validity concerns raised by the trial exclusion rates, the consistent results demonstrated throughout this study across different stimulus sets and participant populations indicate that random response should not have played a major role in affecting experiment results.

Moreover, when considering factors that may have affected the results of these experiments, it is pertinent to consider the role that repeated presentation of stimuli may have had throughout this study. Stimuli were presented four times each within Experiments 2-5 in this thesis. While this is not a method often employed when looking at viewing time and attractiveness, stimuli were repeated in Israel and Strassberg (2009) to evaluate reliability, a foundational paper for the ideas discussed and studied throughout this chapter. In the present study, repetition was utilised to account for the limited data available due to a small participant pool driven by limited funds and time constraints. However, as indicated previously, attractiveness ratings can be influenced by mere exposure (Han et al., 2020; Moreland & Beach, 1992; Rhodes et al., 2005). To minimise the influence of exposure effects, all data points were used, effectively averaging across exposure time. However, it remains possible that repetition influenced attractiveness responses throughout this chapter. In order to address this, an adjustment of methodology would be required whereby each stimulus is presented only once, with an increased participant pool employed. This would be an interesting further avenue of research which would aid to corroborate the findings identified throughout this chapter.

Finally, it should also be noted that all data for the experiments contained within this chapter was collected online. As Experiments 2-5 measured time taken to complete a task, it is important to consider the validity of online data collection for this form of cognitive measurement. When evaluating online latency-based measurement there has been suggestion that while there is promise for measuring time-based response, various factors may have a detrimental effect on the validity of data collection, such as variance in browser use (Chetverikov & Upravitelev, 2016; Crump et al., 2013). However, when experiments are considered which directly compare response measurement online with in-person testing, results suggested that while online testing induces a small delay, this does not add significant noise or influence found effects (Reimers & Stewart, 2007; Schubert et al., 2013). Therefore, results suggest that online data collection can be an effective way of measuring latency-based cognitive response.

Chapter 4

Exploring the Manipulability of Implicit Pupillary Responses of Sexual Interest

Introduction

As explored in Chapters 2 and 3 of this thesis, sexual interest has been investigated through a range of measures including explicit self-report (Dombert et al., 2016; Hoon et al., 1976; Howell et al., 1987) and implicit cognitive latency-based tasks (Ciardha & Gormley, 2009; Israel & Strassberg, 2009; Quinsey et al., 1996). These measures reflect sexual interest in regards to mate preference, which can be categorised as attraction towards, but not within, a group of others. Accordingly, implicit cognitive measures, such as viewing time and Stroop tasks, displayed sensitivity to sex (i.e., male, female) in Chapter 3 but not to within-person changes, such as whether a target was depicted partially clothed or fully dressed. In this chapter a different implicit measure of sexual interest is investigated, which reflects the autonomic measure of pupillary response.

Originally observed in the 1960's, evidence indicates that participants display larger pupils when viewing images congruent with their sexual orientation (Hess et al., 1965; Hess & Polt, 1960). However, initial research aiming to replicate these findings was infrequent and limited due to the rudimentary technology available (Aboyoun & Dabbs, 1998; Green et al., 1979; Hamel, 1974). Within the last decade, contemporary research investigating pupillary responses of sexual interest has increased in popularity (Attard-Johnson & Bindemann, 2017; Rieger et al., 2015; Snowden et al., 2019), due to the development of more reliable technology. In general, these studies have produced corroborating evidence suggesting that larger pupil size occurs when viewing images of sexually interesting others than those of the non-preferred sex (Attard-Johnson et al., 2021; Hess et al., 1965; Rieger & Savin-Williams, 2012).

Pupil size is governed by the parasympathetic and sympathetic branches of the autonomic nervous system (McDougal & Gamlin, 2015; Zele & Gamlin, 2020), and thus occurs automatically and without conscious thought. The primary role of this response is to regulate the pupil light reflex (Ellis, 1981; Guillon et al., 2016; Kun et al., 2012). However, pupil size changes have been observed in connection with a multitude of factors, such as cognitive load (Gavas et al., 2017; Vogels et al., 2018; Zekveld & Kramer, 2014), emotional affect (Bradley et al., 2008; Janisse, 1973; Partala & Surakka, 2003) and sexual arousal (Rieger et al., 2015, 2016). As this response occurs autonomically, it holds promise as a measure potentially impervious to intentional manipulation. This is particularly poignant when considering measurement of socially sensitive factors such as sexual interest, of which some participants may be motivated to conceal genuine response.

However, the extent to which this measure remains unaffected by manipulation remains largely unexplored. When considering pupil size separately from the measurement of a secondary factor, there has been some evidence suggesting that top-down control can affect this response (Ekman et al., 2008; Laeng & Sulutvedt, 2014). In an experiment in which participants were instructed to imagine positive or negative feelings, over half of observers were able to produce larger pupil dilation than in a baseline condition (Ehlers et al., 2015). This was supported by a follow up experiment (Ehlers et al., 2016), which showed that this simple cognitive technique led to enhanced general autonomic arousal, as evidenced by increased skin conductance response alongside enlarged pupil diameter. While these findings were discovered when considering pupillary response in isolation, and not while measuring additional factors, the question arises of whether participants may be able to adopt these techniques while measuring sexual interest.

Another finding which lends support to the potential manipulability of the pupil response comes from the wider attention-based literature. In a task where participants were required to categorise city and mountain images through a button press response, a linear relationship was identified between temporal derivatives of pupil diameter, referring to the extent to which the pupil of the eye constricts and dilates within a specific time window, and task performance (Brink et al., 2016). As such, pupil size was found to decrease in response to a reduction in attention, as indicated through diminished task performance. Corresponding evidence has been observed when considering attentional lapses occurring due to mind-wandering during a breath counting task, whereby pupil size was smaller when participants experienced a period of mind-wandering than when they were task focused (Grandchamp et al., 2014). Taken together, these experiments indicate that the shifting of attention from task-relevant to irrelevant content can affect pupil diameter. However, similarly to the link between pupil size and imagining of feelings, this phenomenon has not been investigated alongside secondary factors, such as sexual interest. There are many techniques which a participant may employ to attempt to manipulate pupil size that have not been adequately investigated, such as averting attention away from a presented stimulus. As such, it remains unknown whether pupillary responses can be controlled through use of a range of techniques or whether such manipulation is sensitive only to specific approaches.

While there has been some evidence suggesting that pupil size may be sensitive to conscious manipulation, if judgements of sexual interest occur automatically, it remains plausible that these may be measurable before conscious control can be exerted. Judgements of attractiveness take place within a fraction of a second (Kaiser & Nyga, 2020; South Palomares & Young, 2018; Willis & Todorov,

2006) and are mandatory (Ritchie et al., 2017), highlighting the automatic nature of assessments of attraction, and thus sexual interest. Furthermore, sexually interesting imagery can engage attention even when presented unconsciously (Jiang et al., 2006). For example, when presented with erotic images that were shown too briefly to be consciously perceived, participants shifted attention to these images instead of control images that were matched for low-level visual content. These effects were congruent with sexual orientation, demonstrating that the sex of the target person had been perceived. As attention can be captured by sexually interesting images, even when these are unconsciously perceived, it remains possible that such findings can be reflected through autonomic responses, such as pupil dilation. Therefore, it seems plausible that pupillary responses of sexual interest may be observable before the influence of any conscious manipulation can be exerted.

Furthermore, there has been some evidence showing that pupil size is impervious to conscious control, when considering pupillary responses separately from sexual interest. Heaver and Hutton (2011) discovered that when assessing familiarity, if participants were instructed to feign amnesia, pupil size still reflected actual familiarity with stimuli, despite attempts to falsify response. While participants were not explicitly asked to control their pupil size, and rather to alter reports of familiarity, results suggest that pupil size is impervious to intentional control, especially when participants are blind to actual measurement method.

While there has been conflicting evidence suggesting whether participants can exert intentional control over their pupillary response, this may not be the only method of manipulating such responses during the measurement of sexual interest. In an experiment investigating the influence of experimenter characteristics on pupillary responses of sexual interest (Chapman et al., 1969), different pupil sizes were elicited dependent on experimenter type. When male participants were in the presence of a formal, business-like experimenter, pupillary responses were comparable across male and female targets. However, a more casual and outgoing experimenter resulted in larger pupil dilation when viewing images of women than men. These person descriptors were informed based on a consensus provided by observers and experimenters and followed a consistent instructional script when presenting the experiment. Therefore, the results from this study suggest that factors external to the observer may influence pupillary responses. This may be particularly exaggerated when examining sexual interest due to the socially sensitive nature of this response. A formal experimenter may cause a participant to be uncomfortable viewing sexually provocative imagery, thus potentially affecting pupil size. However, while this paper provided a basis for understanding the role of different experimenters on pupillary responses of sexual interest, it currently remains unknown how such responses would be affected by the presence of an experimenter *per se*. It may be that *any* pupillary responses to sexually interesting content may be reduced in the presence of other people, due to potential social desirability effects or the perceived social impropriety of such content.

The role of both stimulus-driven attention-based manipulation and external experimenter-driven effects on pupillary responses of sexual interest will be explored throughout the two experiments presented in this study. In Experiment 6, attention-based control will be explored by instructing participants to fixate different areas of a display screen, to either directly focus on or avoid male and female targets. In Experiment 7, the influence of the experimenter will be investigated by analysing pupil size with an experimenter both present and absent from the testing room. Taken together, these findings will provide a strong basis for understanding whether these factors can exert influence over pupillary responses of sexual interest.

Experiment 6

Pupil size is controlled by the sympathetic and parasympathetic branches of the autonomic nervous system (McDougal & Gamlin, 2015; Zele & Gamlin, 2020). The primary function of the pupil response is to regulate the amount of light which reaches the retina (Ellis, 1981; Guillon et al., 2016; Kun et al., 2012). However, pupil size change has been observed in response to a number of factors, including sexual interest (Attard-Johnson et al., 2021; Hess et al., 1965; Rieger & Savin-Williams, 2012). In general, within male participants, findings show increased pupil dilation to images congruent with sexual orientation (Attard-Johnson et al., 2017; Rieger et al., 2015; Watts et al., 2017).

It currently remains unknown the extent to which pupillary responses of sexual interest can be intentionally manipulated. Evidence from studies investigating attentional differences and pupil size have suggested that by employing simple cognitive techniques, such as imagining feelings or mind-wandering, that pupillary responses can be influenced (Ehlers et al., 2016, 2015; Grandchamp et al., 2014). However, the potential manipulability of this response has not yet been empirically investigated when exploring factors such as sexual interest.

The aim of this experiment was to explore the potential influence of attentional control when investigating pupillary response of sexual interest, by manipulating where on a display screen participants fixate. This design was implemented to assess the potential impact of a participant attempting to control cognitive response by averting attention away from a stimulus target. To this end, heterosexual male participants were instructed to view one of three areas of a display screen while an image of a male or female was presented. These areas consisted of the person, the background, or a small box presented on the background. Given evidence showing that sexually interesting imagery draws attention even when presented subconsciously (Jiang et al., 2006), it is possible that person images will attract attention and be perceived regardless of where on the screen participants are instructed to view. If pupillary responses of sexual interest are impervious to attentional control, larger pupils should be exhibited during the viewing of female stimuli. This would persist as, although fixation location is constrained, attention will still be drawn towards these female images despite attempts to retain this attention to various areas of the screen. However, if attempts to control attention can influence pupil size, comparable pupillary responses should be observed in the male and female conditions when viewing the background or the box, with pupil diameter reflective of sexual interest still exhibited when viewing the person image.

Method

Participants

Twenty-eight male undergraduates from the University of Kent participated in this experiment in return for course credits (M_{age} = 22.68 years, *SD* = 4.68, Range = 18 to 36). To take part in this experiment, participants were required to score either '0' or '1' on the Kinsey scale (Attard-Johnson et al., 2021; Hess et al., 1965; Rieger & Savin-Williams, 2012), indicating heterosexuality. An additional four participants were tested but excluded as they failed to meet these criteria.

Materials

The stimuli comprised 36 Caucasian models (18 male, 18 female). Model anonymity was ensured by merging two images, resulting in a head and body belonging to different identities. Stimuli resolution was 66pi with a maximum height and width of 21.4 x 12.5cm, with each image covering comparable screen area. Stimuli were modified for the condition in which participants were asked to exclusively view a box presented on the background of the image. In these instances, a 3.5 x 3.5cm square was presented to the left and right of each stimuli model, situated in the vertical centre of the screen, see Figure 4.1 for example images. The nearest side was set at a distance of 10.7cm from the horizontal midpoint of the screen.

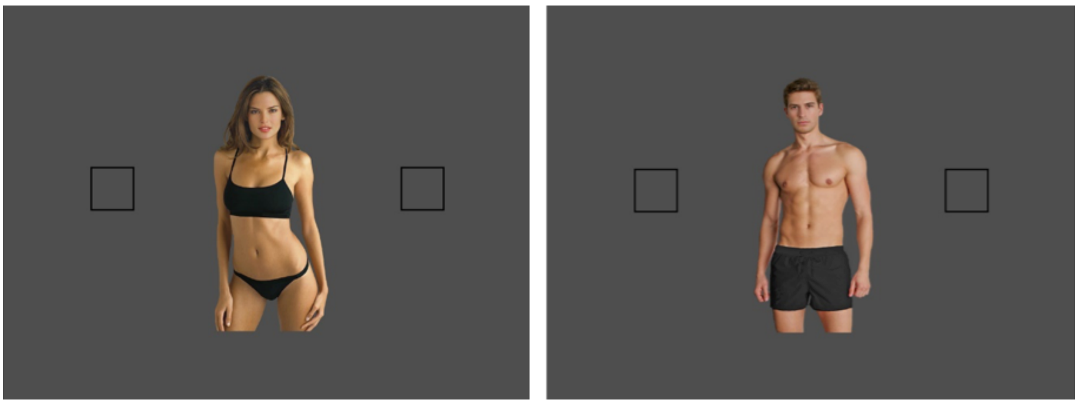

Figure 4.1. Example of stimuli presented during the placeholder condition featuring a male and a female model.

Eye-Tracking Equipment:

A SR-Research EyeLink 1000 eye-tracker was used to record eye movements and pupil size throughout this experiment. The EyeLink 1000 has a gaze position accuracy of less than 0.5° and spatial resolution of less than 0.01° of visual angle. It features a 1,000 Hz sampling rate with pupil size resolution of 0.1%. To compute pupil size, this eye-tracker uses a video-based infrared camera and measures the pixels which are obstructed by the pupil. Both corneal reflection and dark pupil are computed to determine pupil dilation. Pupil size was recorded at each fixation point as an integer between 400 and 16,000. The stimuli were presented on a 21” colour monitor, with a resolution of 1,024 x 768 pixels.

Procedure

This experiment was conducted in a quiet, windowless room with controlled artificial lighting. To facilitate accurate data collection, head movements were controlled using a chinrest positioned 60cm away from the monitor. The built-in auto-threshold function was utilised alongside the standard 9-point calibration process, in which participants attend separately to nine points presented on the screen. This process was then repeated to validate these responses. Calibration was accepted if tracking accuracy was better than 0.5° of visual angle.

This experiment contained three separate blocks in which participants were asked to limit attention to either a person, the background or a box presented on the background. Presentation order of these conditions was counterbalanced across participants, who were informed where to direct their attention before the start of each block. Participants were not informed of the nature of this experiment, beyond brief mention of attractiveness and familiarity being the topic of study in the pre-experiment information form, and thus there was no instruction to hide or limit attraction present throughout this experiment. Proceeding each trial, participants attended to a drift correction point presented randomly on the left or right of the screen. Each block contained 20 trials lasting roughly five minutes. Trials began with a 1-second grey screen, followed by 5-seconds where the stimulus was presented, and concluded with a 1-second grey screen. Participants were offered a short break upon block completion, after which the calibration process was repeated.

Results

Data Preparation

Any fixations which occurred outside of the display screen border, or which were concealed by blinking, were excluded from data analysis. Furthermore, fixations made within 80ms were merged with the previous or subsequent fixation if they were within half a degree of visual angle (for similar approaches, see e.g., Attard-Johnson et al., 2016; Bindemann et al., 2009, 2010). For the purpose of this experiment, initial fixations (that may have not occurred in response to stimuli presentation) were not removed as they were expected to average out across stimuli, and thus would not affect the overall results. This was corroborated by analyses which did remove initial fixation and found no meaningful differences in results. Raw data was utilised for these analyses due to research indicating that this measure is a reliable way of measuring pupil size change (Attard-Johnson et al., 2019). Pupillary responses were calculated by taking a measurement of pupil size at each fixation for each condition, and then averaging this across all participants, resulting in cross-subject means. Analyses were then conducted with this data to investigate fixation locations and pupil size during the viewing of a person image, the background or a box presented on the background. A post-hoc power analysis was calculated for this experiment and indicated a statistical power of .07.

Regions of Interest

To ensure that participants were adhering to task instructions, fixation locations were first analysed. Each image was separated into three regions of interest (ROI) comprising of the body, the background and a box presented on the background (see Figure 4.1). Figure 4.2 represents the percentage of fixations occurring on the target ROI that participants were asked to view in each of the three conditions. For example, for the person condition, these data capture fixations on the male and female targets, whereas the background condition captures fixations on the grey image background. These data show that the majority of fixations were made consistent with the task demands.

A 2 (Sex: male, female) x 3 (Condition: person, background, box) repeated-measures ANOVA indicated that no main effect of Sex, $F(1,27) = 0.47$, $p = .50$, $\eta^2 = .00$, or interaction between Sex and Condition was found, $F(2,54) = 2.91$, $p = .06$, $\eta^2 = .005$. However, a main effect of Condition was identified, $F(2,54) = 11.58$, $p < .001$, $\eta^2 = .20$. Post-hoc *t*-tests, with Sex collapsed across Condition, were computed to investigate this main effect, with a Bonferroni adjusted alpha level of .017 (.05/3). There was no main difference identified between the person and the background conditions, $t(27) = 0.82$, $p = .42$, $d = .15$. However, fewer fixations occurred in the box condition than in the person, $t(27) = 3.39$, $p = .002$, $d = .64$, and the background conditions, $t(27) = 4.07$, $p < .001$, $d = .77$. As the box presented to participants was smaller in size compared to the person and background, it is possible that reduced fixations may be due to natural eye saccades causing fixations to occur outside of the bounds of the box.

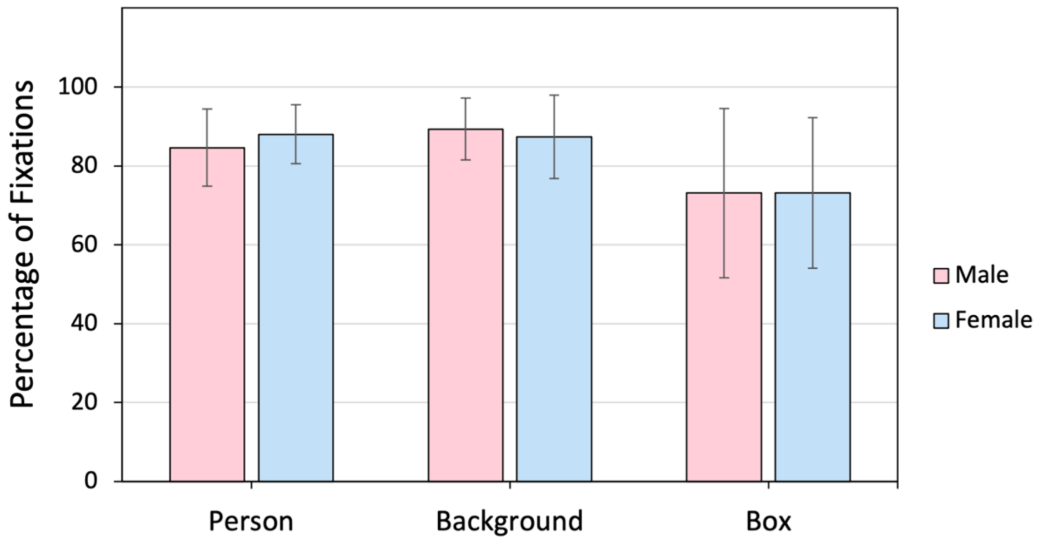

Figure 4.2. Bar graph representing the mean number of fixations made to the instructed ROI per stimulus condition for male and female targets. Error bars represent standard deviation.

To determine whether the presentation of a person image draws attention even when participants are instructed to attend elsewhere on the display screen, the percentage of fixation locations occurring on the person was considered for each condition. As participants generally followed task instructions, there was too little data within the background and box conditions to perform inferential statistics. However, as can be seen in Figure 4.3, the mean number of fixations occurring on the person suggest that images of people can be ignored when observers are asked to view elsewhere on the screen.

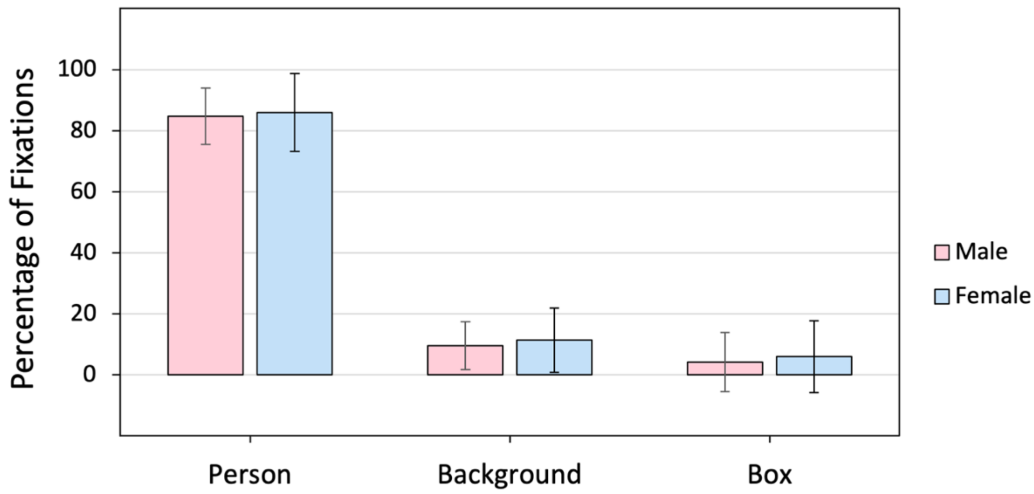

Figure 4.3. Bar graph indicating the mean number of fixations occurring on the person within each condition, with error bars representing standard deviation.

Pupillary Responses

While it was not possible to explore whether a person image effectively drew attention via fixation location due to limited data, the analysis of pupil size can help to assess this question. Pupillary responses were analysed both when occurring anywhere on the screen, and when transpiring specifically on the target ROI. By constraining analysis to specific ROIs this also eliminates any potential influence of luminance by removing any noise in pupillary responses attributable to brightness differences between the person and the background. If images of people that are congruent with the gender of one's sexual interest elicit higher pupil dilation, and these targets draw attention regardless of where on a screen an observer is instructed to view, then these analyses should indicate sexual interest when fixation location is unrestrained, but not when considered solely in relation to the instructed ROI.

Firstly, pupil size was analysed based on all fixations for each ROI, regardless of where participants were looking on the display, see Figure 4.4 for mean values. A 2 (Sex: male, female) x 3 (Condition: person, background, box) repeated-measures ANOVA was then computed to investigate whether pupillary responses are affected by the sex of a stimulus target or location of attention within an image. No main effect of Sex, $F(1,27) = 1.18$, $p = .29$, $\eta^2 = .00$, or interaction between Sex and Condition was found, $F(2,54) = 0.85$, $p = .43$, $\eta^2 = .00$. However, a main effect of Condition was identified, $F(2,54) = 3.50$, $p = .04$, $\eta^2 =.01$. Post-hoc *t*-tests, with Sex collapsed across Condition, with a Bonferroni alpha adjustment of 0.017 (.05/3) indicated marginally larger pupil diameter during the background than the person condition, $t(27) = 2.49$, $p = .019$, $d = .47$. No differences were found between the person and box, $t(27) = 2.07$, $p = .05$, $d = .39$ or the background and box conditions, $t(27) = 0.80$, $p = .43$, $d = .15$.

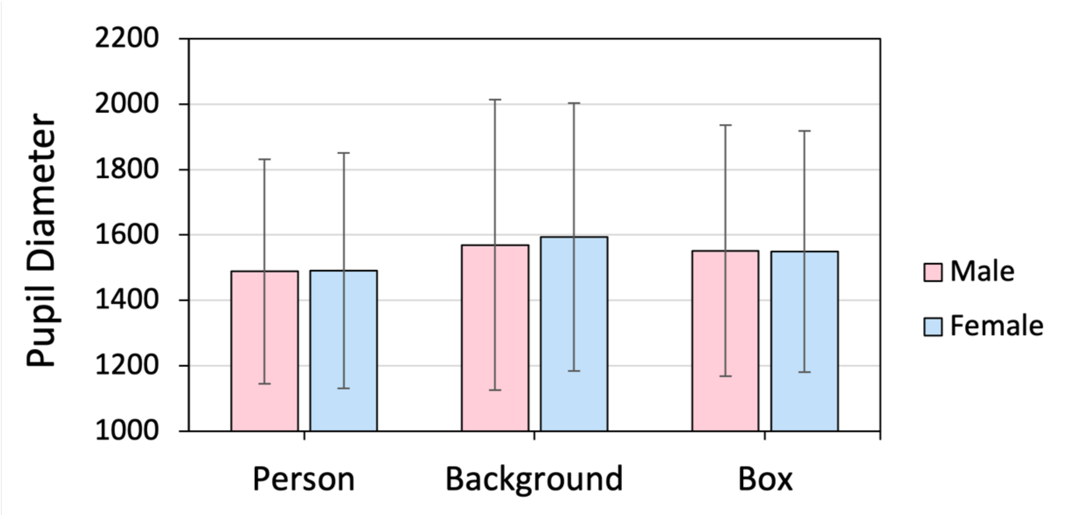

Figure 4.4. Bar graph depicting mean pupil size viewing the person, the background and the box, split by male and female targets. Error bars represent standard deviation.

To investigate whether pupillary response was reflective of where participants were actually fixating, and not the potential influence of the person target, this analysis was repeated with any fixations occurring outside of instructed ROI excluded. See Figure 4.5 for mean pupil size across Sex and Condition when fixation location was constrained consistent with the condition instructions. The 2 (Sex: male, female) x 3 (Condition: person, background, box) repeated-measures ANOVA showed no main effect of Sex, $F(1,27) = 1.92$, $p = .18$, $\eta^2 = .00$, or interaction between Sex and Condition, $F(2,54) = 0.86$, $p = .43$, $\eta^2 = .43$. A main effect of Condition was found, $F(2,54) = 3.74$, $p = .03$, $\eta^2 = .01$. Post-hoc *t*-tests, with Sex collapsed across Condition, were computed with a Bonferroni adjusted alpha level of .017 (.05/3) and identified that observers elicited marginally smaller pupil sizes when viewing a person than the background, $t(27) = 2.50$, $p = .02$, $d = .47$, or the box presented on the background, $t(27) = 2.53$, $p = .02$, $d = .48$. There was no difference between the background and box conditions, $t(27) = 0.37$, $p = .72$, $\eta^2 = .07$. As these results replicate those found when fixation location was unconstrained, it appears likely that the person target did not draw visual attention when participants were instructed to view elsewhere on the screen.

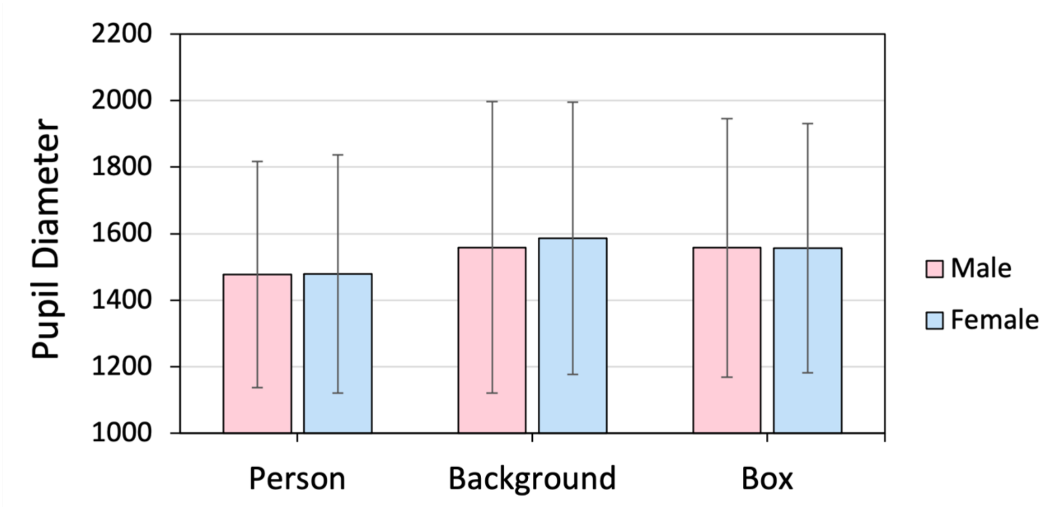

Figure 4.5. Bar graph showing mean pupil size when viewing either an image of a person, a background or a box presented on a background, separated for male and female targets. Error bars indicate standard deviation.

Discussion

The aim of this experiment was to investigate whether pupil dilation occurs during the viewing of individuals of an observers' preferred sex when their focus of attention is intentionally manipulated. Participants were asked to restrict their eye gaze to either a person, the image background, or a box positioned on the background. Fixation locations were analysed as an indicator that participants were adhering to task instructions, with the instructed ROI generally being viewed over 80% of the time. A notable exception was the box condition, in which 73% of the fixations occurred on the box in comparison with the rest of the display. These differences may be accounted for by the small size of the presented box, as it is possible that some fixations were unintentionally located outside of the box perimeter. In these cases, fixations would have been calculated towards the background rather than the box. The high rate of background fixations identified within this condition (23.89% for male and 16.56% for female targets) support this suggestion. Comparable fixations to ROI's across conditions, when accounting for increased background viewing in the box condition, show that participants were accurately following task instructions. As such, it can be concluded that differences in viewing pattern across conditions cannot account for any differences in pupil size.

It has been widely documented that heterosexual male participants exhibit larger pupil dilation to images of women than men (Attard-Johnson et al., 2017; Hess et al., 1965; Rieger & Savin-Williams, 2012). However, in this experiment, pupil size was comparable regardless of stimulus sex. As fixation focus did not interact with pupillary responses of sexual interest, this initially suggests that pupil size cannot be wilfully influenced by looking strategies and differences in attention-based behaviour.

However, comparable pupil sizes were also found when considering the person target block in isolation, indicating that sexual interest did not drive pupillary responses *throughout* this experiment. As explored, when investigating sexual interest without an attentional manipulation, pupil size generally indicates sexual orientation (see, e.g., Attard-Johnson et al., 2021; Attard-Johnson & Bindemann, 2017; Rieger et al., 2015). When introducing an attention-based manipulation in this experiment, through constraint of fixation locations, these effects were eliminated. This therefore suggests that pupillary responses may have been reflective of attention and not sexual interest throughout this experiment, a factor which has been identified alongside pupil size change when considered in isolation from sexual interest (Alnæs et al., 2014; Kang et al., 2014; Smallwood et al., 2011). As such, future experiments investigating sexual interest through this measure should be careful to consider the role of attention, as this factor may obscure potential results if not carefully controlled. This is considered further in the General Discussion of this chapter.

Finally, pupil size was consistently higher for the background than the person throughout this experiment, which may be accounted for by the lower luminance present in comparison with the person images. While this factor may have exerted influence over pupillary responses generally, target sex should still effect pupil size if driven by sexual interest. As such, the difference in luminance between the person image and the background is not of primary concern in this experiment as sexual interest responses should still be expressed when looking at the person, or the background, separately.

Experiment 7

Experiment 6 provided a basis for understanding whether pupil size can be consciously manipulated while measuring sexual interest. Experiment 7 was designed to expand these findings and to explore the potential influence of external experimenter-based control over pupillary responses of sexual interest. Prior evidence has shown that an experimenter's behavioural characteristics can exert influence over pupil size (Chapman et al., 1969). However, it remains unknown whether the presence or absence of an experimenter can also impact pupil dilation. This experiment will assess whether pupillary responses of sexual interest are influenced by whether or not the experimenter is in the room at time of testing.

To examine this, a slightly different approach was devised. The stimuli now comprised of magazine covers featuring partially clothed men and women as they provided both a person target to induce attraction and an interesting background for participants to attend to if they feel motivated to avoid the target. Eye-gaze was investigated alongside pupil size in this experiment to identify if the presence of an experimenter effects pupillary response, and whether potential differences in pupil size across experimenter conditions can be accounted for by variance in fixation patterns within a stimulus. If the presence of an experimenter affects pupillary responses of sexual interest, then we would expect to find differences in pupil size when viewing male and female images depending on whether the experimenter was present or absent during experiment completion.

Method

Participants

Thirty-four male participants took part in this experiment (M_{age} = 22.54 years, *SD* = 4.86, Range = 18 to 36). Participants were the same as those who took part in Experiment 6, with an addition of two participants included to combat the potential experimental errors that arose. All participants self-identified as heterosexual according to the Kinsey scale (Kinsey et al., 1948), with the exception of 3 participants, who were consequently excluded from analysis. A further two participants were excluded from data analysis due to complications with the experimental session, resulting in the analysis of 29 data sets.

Materials

The stimuli utilised in this experiment were 28 covers taken from Men's Health and Women's Health magazines. Half of these covers feature men and half women, each of which were partially clothed in athleisure wear. Example stimuli are shown in Figure 4.6. The magazine cover measured a height of 551 and width of 402 pixels, with a pixel per inch resolution of 72. All stimuli targets covered a comparable area of the magazine. These covers were then superimposed onto a grey background.

Figure 4.6. Example of the magazine stimuli presented in Experiment 7. The left image is taken from Women's Health and features a female stimulus model. The right image features a male model from the magazine Men's Health.

Procedure

This experiment took place in the same controlled room and utilised the same equipment as Experiment 6, and the same calibration process and stimulus presentation time frame was followed. To ensure any potential experimenter effects remained consistent, all testing sessions were run by the same experimenter.

This experiment consisted of three blocks. The first two blocks included a free-viewing task, whereby participants were instructed to view the stimuli however they wanted, with no responses required. In one of these blocks, the experimenter was present in the room as participants performed this task. In the other block, the experimenter was absent. In this condition, the experimenter excused themselves by clarifying that for the current block they would not be in the room and providing instructions for how to begin the experiment once absent. The order of the experimenter present and absent conditions was counterbalanced across participants. In a final block, participants were presented with all stimuli shown previously in blocks 1 and 2, and asked to rate the attractiveness of the model on a scale from 1 (not attractive) to 7 (highly attractive).

Results

Data Preparation

The data were processed as in Experiment 6. Thus, any fixations which occurred outside of the display screen border, or which were concealed by blinking, were excluded from data analysis. Fixations made within 80ms were merged with the previous or subsequent fixation if they were within half a degree of visual angle (for similar approaches, see e.g., Attard-Johnson et al., 2016; Bindemann et al., 2009, 2010). In line with the procedure utilised in Experiment 6, initial fixations were not excluded as their impact does not meaningfully effect the results. Raw data was utilised as research suggests that this measure is a reliable method in the analysis of pupillary responses (Attard-Johnson et al., 2019). Pupillary responses were calculated by taking a measurement of pupil size at each fixation for each condition, and then averaging this across all participants, resulting in cross-subject means. A post-hoc power analysis indicated that this experiment had a statistical power of .06.

Ratings

Firstly, attractiveness ratings were investigated to ascertain whether participants, all of whom self-identified as heterosexual men, considered female ($M = 5.38$, $SD = 0.67$) targets to be more attractive than male ($M = 3.07$, $SD = 1.52$). This was corroborated by a pairwise *t*-test which indicated that women were rated as more attractive than men, $t(28) = 8.50$, $p < .001$, $d = 1.58$.

Regions of Interest

To assess whether eye gaze was affected by the sex of the target images or experimenter presence, the number of fixations to two ROIs were computed, which comprised of the Person and the Magazine background. The mean number of fixations to these ROIs are illustrated in Figure 4.7).

To analyse these data, two repeated-measures ANOVAs were computed, one for each ROI. Firstly, fixations on the Person were investigated. A 2 (Sex: male, female) x 2 (Experimenter: present, absent) repeated-measures ANOVA showed no interaction between Sex and Experimenter, $F(1,57) = 0.16$, $p = .69$, $\eta^2 = .00$. Furthermore, no main effect of Experimenter was identified, $F(1,57) = 0.13$, $p = .72$, $\eta^2 = .00$. However, a main effect of Sex was found, $F(1,57) = 30.27$, $p < .001$, $\eta^2 = .07$ with more fixations recorded on the target persons in female than male images (see Figure 8).

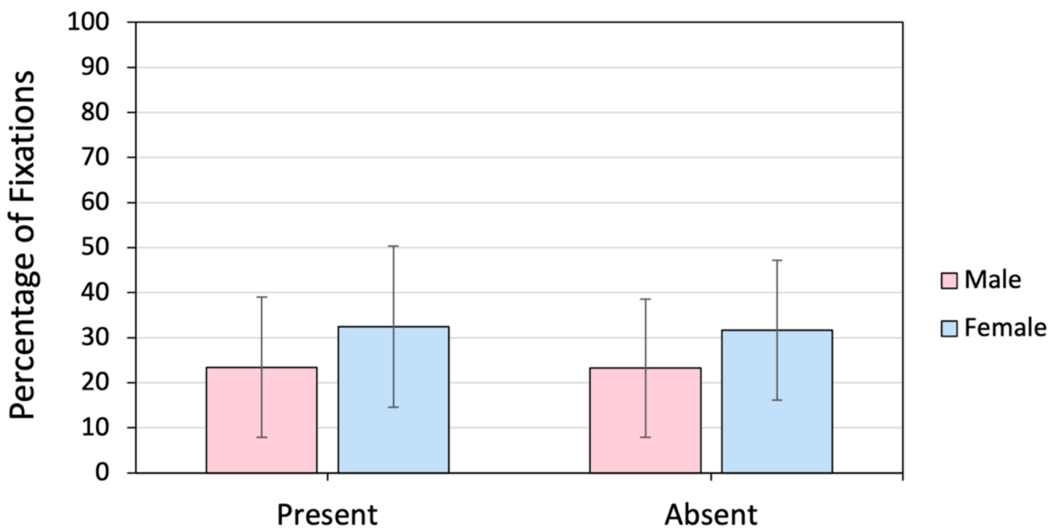

Figure 4.7. Bar graph representing average percentage of fixations on the person target per target sex and experimenter presence. Standard deviation is indicated through use of error bars.

This analysis was then repeated for the fixations occurring on the magazine background. A 2 (Sex: male, female) x 2 (Experimenter: present, absent) repeated-measures ANOVA again indicated no interaction between Sex and Experimenter, $F(1,28) = 0.01$, $p = .92$, $\eta^2 = .00$, or main effect of Experimenter, $F(1,28) = 0.001$, $p = .97$, $\eta^2 = .00$. A main effect of Sex was again found, $F(1,28) = 37.82$, $p < .001$, $\eta^2 = .16$, which was accounted for by more fixations on the magazine backgrounds in the male than the female target condition (see Figure 4.8).

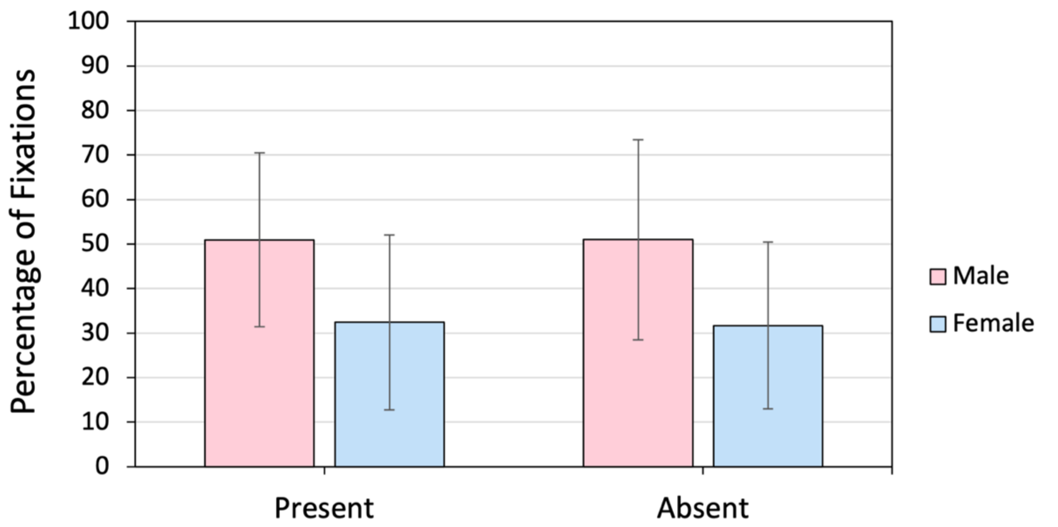

Figure 4.8. Bar graph showing mean fixations to the magazine background by target sex and experimenter presence. Error bars represent standard deviation.

Pupillary Responses

Pupil size was then investigated to explore whether the sex of the stimuli or presence of the experimenter affected pupil diameter. These analyses were again split by fixations to the Person or Magazine cover.

For fixations on the person ROIs, a (Sex: male, female) x 2 (Experimenter: present, absent) repeated-measures ANOVA indicated a main effect of Sex, $F(1,57) = 8.04$, $p = .006$, $\eta^2 = .001$, due to larger pupil size when viewing male than female targets (see Figure 4.9). Furthermore, a main effect of Experimenter was identified, $F(1,57) = 9.45$, $p = .003$, $\eta^2 = .004$, with larger pupil diameter recorded when the Experimenter was present, than when absent. An interaction was also found between the factors of Sex and Experimenter, $F(1,57) = 4.84$, $p = .03$, $\eta^2 = .001$. Post-hoc pairwise *t*-tests with a Bonferroni adjusted alpha level of .0125 (.05/4) identified that viewing male targets led to larger pupil size than female targets when the Experimenter was absent from the room, $t(57) = 3.59$, $p < .001$, $d = .47$. Furthermore, pupil dilation was larger when observing female targets when the Experimenter was present rather than absent, $t(57) = 3.63$, $p < .001$, $d = .48$. There was no difference in pupil size when considering male and female targets when the Experimenter was present, $t(57) = 0.41$, $p = .68$, $d = .05$, or when observing male targets when the Experimenter was present and absent, $t(57) = 1.82$, $p = .08$, $d = .24$.

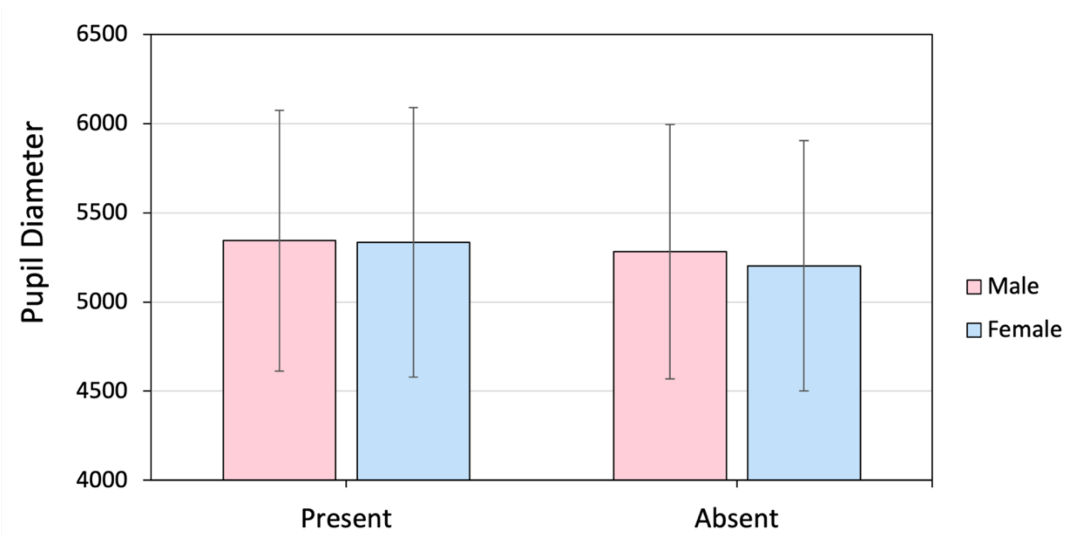

Figure 4.9. Bar graph depicting mean pupil size while viewing a person target for target sex and experimenter presence. Error bars indicate standard deviation.

Lastly, pupil size was considered for fixations occurring on the magazine background. Figure 4.10 indicates mean pupil diameter and shows comparable results across Sex and Experimenter. A (Sex: male, female) x 2 (Experimenter:

present, absent) repeated-measures ANOVA was conducted to explore these findings. No main effects of Sex, $F(1,27) = 0.32$, $p = .58$, $\eta^2 = .00$, or Experimenter were found, $F(1,27) = 1.89$, $p = .18$, $\eta^2 = .002$. There was an interaction found between these factors, $F(1,27) = 5.17$, $p = .03$, $\eta^2 = .001$. However, when pairwise *t*-tests with a Bonferroni adjusted alpha level of 0.0125 (0.05/4) were computed, no significant relationship was identified for male targets when the Experimenter was present and absent, $t(28) = 0.24$, $p = .81$, $\eta^2 = .05$, male and female targets when the Experimenter was present, $t(27) = 1.35$, $p = .19$, $\eta^2 = .26$, male and female targets when the Experimenter was absent, $t(28) = 1.94$, $p = .06$, $\eta^2 = .36$, or female targets when the Experimenter was present and absent, $t(27) = 2.25$, $p = .03$, $\eta^2 =.43$.

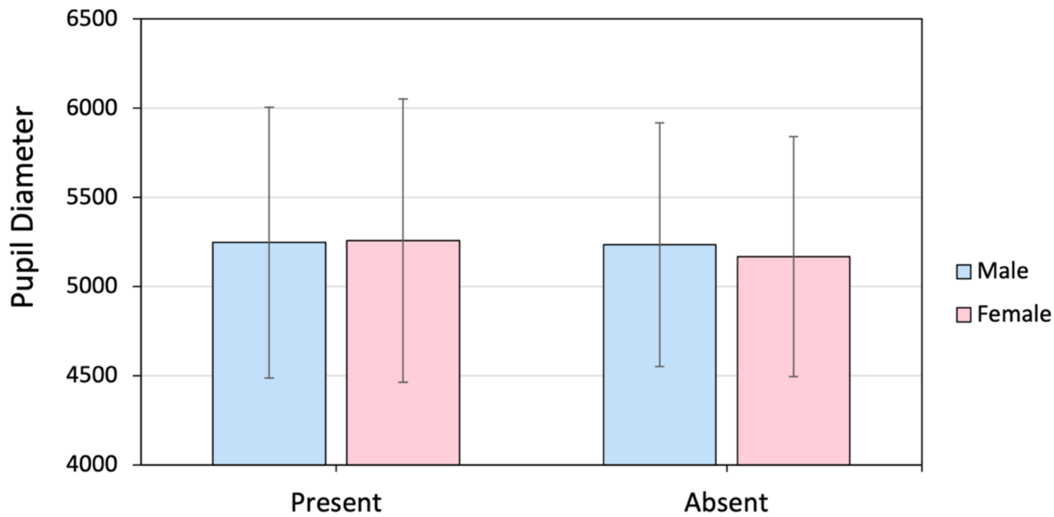

Figure 4.10. Bar graph illustrating mean pupil diameter per target sex and experimenter presence, when considering fixations made to the magazine background. Error bars represent standard deviation.

Discussion

The aim of Experiment 7 was to investigate whether pupillary responses of sexual interest can be influenced by whether an experimenter is present or absent during testing. Attractiveness ratings were first explored to ensure that the heterosexual male participants considered female targets to be more attractive than male targets. This assumption was supported by the ratings data, indicating that explicit responses can effectively measure sexual orientation. This suggests that the stimuli presented were able to promote feelings of sexual interest, allowing pupil size to be analysed accordingly.

Participants viewing patterns were then investigated to explore whether observers fixated on images differently across target sex or experimenter presence. Viewing patterns were comparable both when the experimenter was present or absent, thus ensuring that any pupil size differences within this factor cannot be accounted for by viewing pattern. However, participants did show different fixation patterns across target sex, with female targets receiving more fixations than male targets. In turn, the magazine background of male targets received more fixations than the background of female targets. As all participants in this experiment self-identified as heterosexual males, these viewing patterns can be considered to reflect sexual interest, as participants looked longer at female targets and actively avoided observing the male targets. However, an alternative explanation for this phenomenon may be that the magazine backgrounds for the male stimuli were simply more entertaining than that of their female counterparts. In order to keep the stimuli in this experiment as generalisable to real life as possible, real life magazine covers were presented *without* any content edits. However, a replication of this experiment, with the magazine backgrounds directly comparable across stimulus model sex, could be an effect way of analysing what drove the target sex viewing pattern differences in this experiment.

Lastly, pupillary responses were considered to explore whether pupil size reflecting sexual interest can be controlled by experimenter presence or absence at time of testing. Firstly, this analysis did not reveal pupillary responses reflective of sexual interest, with pupil size remaining comparable when viewing the magazine background regardless of the target sex. When considering pupil size when directly viewing the target, larger dilation occurred for male than female targets, a surprising finding that diverges from those found within the general literature (Attard-Johnson et al., 2021; Ó Ciardha et al., 2018; Rieger et al., 2015). However, pupil size across this experiment was generally high, especially in comparison with Experiment 6. For example, when considering all pupil sizes in both Experiment 6 and 7, a paired T-test indicated that pupil diameter was higher throughout Experiment 7, $t(191) = 61.88$, $p < .001$, $d = 4.47$. This suggests that the stimuli in Experiment 7 may have been particularly arousing. Increased pupil dilation has been identified in response to high levels of arousal when investigating sexual interest (Snowden et al., 2019). In an experiment where participants were presented with clothed and nude images, the higher arousing nude imagery led to increased pupil size regardless of the sex of the observer or the target (Aboyoun & Dabbs, 1998). Therefore, it is possible that results pertaining to sexual interest in this experiment were masked by a strong general arousal response.

While pupillary responses of sexual interest were not exhibited in this experiment, the presence of an experimenter did affect this measure. Larger pupil sizes were exhibited when viewing the person target, whilst the experimenter was present in the room. This may be due to a general period of increased attention promoted by experimenter presence, whereby participants are more motivated to pay attention to the experiment when they know their responses are being evaluated. This increase in attention has been documented alongside pupil size in the wider literature, as evidence suggests that larger pupil dilation occurs as attentional effort increases (Alnæs et al., 2014; Kang et al., 2014; Karatekin et al., 2004).

General Discussion

This chapter presents two experiments, which examined whether pupillary responses of sexual interest can be controlled. Experiment 6 investigated whether pupil size can be manipulated by intentionally directing attention away from a person target by constraining fixation locations to specific screen areas. In this experiment, pupil size was comparable when viewing a person, the background or a placeholder box presented on the display background. The absence of pupil dilation differences between male and female images when viewing the person targets suggests that pupil size was not reflective of sexual interest. Experiment 7 explored whether pupillary responses can be manipulated by a force external to the participant, such as the presence or absence of an examiner. Despite different viewing patterns for male and female targets that corresponded with observers' sexual interest, pupillary responses were not indicative of sexual orientation in this experiment. However, it should be noted that the statistical power of these experiments were very low, as indicated by post-hoc analysis. As such, the results need to be considered in relation to such low power. In order to address this, further testing with a larger participant sample is recommended.

These experiments highlight the potential influence of attention on pupillary responses. Despite consistent evidence that heterosexual male observers exhibit larger pupil dilation to images of women than men (Attard-Johnson et al., 2021; Hess et al., 1965; Rieger & Savin-Williams, 2012), this pattern was not present within this study. These results occurred despite explicit rating data reflecting sexual orientation. As such, results suggest that pupil size was sensitive to a separate factor, rather than being driven by sexual interest. The consistent secondary factor manipulated throughout these experiments was attention. In Experiment 6, attention was controlled by instructing participants to view specific areas of a stimuli, while Experiment 7 explored this factor through presence of an experimenter.

The role of attention in influencing pupil size has been widely documented outside of the sexual interest literature, with increased attentional effort leading to enhanced pupil size (Alnæs et al., 2014; Karatekin et al., 2004; Smallwood et al., 2011). For example, in an experiment where participants were presented with digit sequences and asked to either indicate or recall whether a number was odd or even, increased pupil dilation was found to synchronise with stimulus presentation when attending to the task than when mind-wandering (Kang et al., 2014). Furthermore, evidence suggests that the superior colliculus modulates pupil size changes to salient stimuli (Joshi & Gold, 2020). This is a brain area which has been consistently identified to contribute to the control of attention (Ignaschchenkova et al., 2004; Müller et al., 2005; Schneider & Kastner, 2009), thus suggesting that pupil dilation can be influenced by attention. Due to this, when considering the tasks employed in both experiments within this chapter, the role of attentional effort in influencing pupil size may account for the non-discriminant responses identified for sexual interest. Both experiments employed tasks in which the attentional effort requirement may have been high, either due to specific instructions restricting eye-gaze or through increased pressure to attend to the task due to experimenter presence. As such, it is plausible that pupil size was reflective of attentional effort effects rather than sexual interest.

The likelihood of pupillary responses relating to attentional effort rather than sexual interest within this study is supported by the salience of these factors within the experimental tasks. While participants were instructed where to view, or of the experimenter's presence within the room, no explicit mention was made in reference to the sex of the targets. The only potential prompt relational to sexual interest was made through a pre-testing demographic form, whereby participants were asked to indicate their sexual orientation. Thus, the salience of attention-based cues may have led to resulting pupillary responses. However, research suggests that impressions of attractiveness occur quickly (Locher et al., 1993; South Palomares & Young, 2018; Willis & Todorov, 2006) and automatically (Ritchie et al., 2017), with sexually attractive images drawing attention even when presented subconsciously (Jiang et al., 2006). As such, it is likely that sexual interest was considered by participants throughout the current study despite the higher salience of attention-based instructions. Therefore, it may be that the influence of attentional effort masked sexual interest related responses, rather than stopping these judgements from occurring altogether.

When considering the autonomic nature of attractiveness judgements, it is pertinent to explore the attentional demands required to make such assessments. When considering spontaneous trait inferences, evidence suggests that trait inferences are dependent on memory (Wells et al., 2011; Zhou et al., 2021). For example, when participants are presented with photographs and descriptions of trait-implying behaviours and asked to rate said traits, spontaneous trait inference depends on working memory capacity (Wells et al., 2011). As such, for these trait judgements to occur, working memory must be employed. Similarly, in the context of Experiment 6, the primary task demands of attention direction may also have required working memory. Evidence suggests that directing attention and working memory are highly interrelated concepts (Kiyonaga & Egner, 2013; Oberauer, 2019; Soto et al., 2005). This is evident from interference paradigms. For example, when participants are instructed to classify a directional arrow when this is flanked by congruent or incongruent directional arrows, participants were less accurate when working memory load is also high (Pratt et al., 2011). This phenomenon suggests that working memory is required for accurately directing attention. As both the process of directing attention and performing trait inferences require working memory, and Experiment 6 explicitly required participants to attend to specific display locations, it is possible that the explicit task requirements left insufficient working memory capacity for effective attractiveness trait judgements. However, working memory load was not directly investigated within this experiment, therefore its potential interactive effect remains inconclusive.

While the primary task demand of directing attention may have diminished pupillary responses of sexual interest in Experiment 6, Experiment 7 did not include such demands. Instead, this Experiment allowed participants to freely view magazine images featuring partially clothed models. Akin to Experiment 6, no pupil size differences reflective of sexual orientation were identified. In this instance, it is plausible that a general increase in arousal promoted by the presentation of sexually laden stimuli occurred. Increases in pupil dilation have been observed concurrently with sexual arousal (Bernick et al., 1971; Rieger et al., 2015). While some evidence has suggested that the level of stimuli explicitness does not (Attard-Johnson & Bindemann, 2017), or to a small effect can (Watts et al., 2017), influence pupillary responses of sexual interest, contradictory evidence also suggests that the high level of arousal promoted by sexually appealing imagery may mask such responses

(Snowden et al., 2019). For example, in an experiment which presented observers with nude and clothed targets, larger pupil dilation was recorded when viewing nude imagery regardless of target sex (Aboyoun & Dabbs, 1998).

The stimuli in Experiment 7 were partially clothed and thus sexually arousing, as evidenced by the large average pupils observed towards all images. As such, it seems plausible that the increased sexual arousal elicited by all stimuli within this experiment may have diminished any pupil size differences due to target sex. Furthermore, the magazine stimuli featured non-sexual elements, such as written headlines and slogans. These additional features may have also elicited more general types of arousal. Emotional arousal, for example, can also elicit pupil dilation (Bradley et al., 2008; Janisse, 1973; Wang et al., 2018) and may have been triggered by some of the written stimulus content. Therefore, it is possible that arousal promoted by the non-person stimulus materials presented within this experiment might account for the absent pupil size effects between sex categories.

In conclusion, this study constituted the first empirical investigation aiming to investigate whether pupillary responses of sexual interest can be intentionally manipulated. The results suggest that pupil size can be affected by task demands, such as the intentional direction of attention through the constraint of visual fixation locations, resulting in comparable pupil dilation regardless of target sex. As such, future investigations of sexual interest through pupillary responses should be conscious of the level of attentional effort demanded through experimental design. Furthermore, the results also indicated that highly arousing stimuli can enlarge pupils regardless of sexual orientation, as evidenced by the particularly large pupil sizes observed in Experiment 7. Taken together, this chapter suggests that the implicit pupil response measure of sexual interest may not be as robust as previously thought, as potential effects can be eliminated by task demands and high arousal.

Chapter 5

Summary and Discussion

This thesis investigated the sensitivity of cognitive measures of sexual interest and familiarity, and explored the extent to which these can be manipulated. The introduction reviewed the current measures employed in the evaluation of sexual interest, with consideration given to discussing their merits and limitations. These measures range from explicit tasks whereby sexual interest is garnered through observer self-report (Dombert et al., 2016; Hoon et al., 1976; Howell et al., 1987) to more implicit forms of measurement. These included cognitive latency-based measures of viewing time (Ebsworth & Lalumière, 2012; Imhoff et al., 2010; Mokros et al., 2013) and Stroop tasks (Ciardha & Gormley, 2012; Mannfolk et al., 2023; Ó Ciardha & Gormley, 2013) as well as paradigms aiming to explore sexual interest through arousal of the autonomic nervous system, such as genital (Barbaree et al., 1979; Barker & Howell, 1992; Müller et al., 2014) and pupillary responses (Attard-Johnson et al., 2017; Hess et al., 1965; Rieger et al., 2015).

When considering the utility of these paradigms for the measurement of sexual interest, several potential caveats had yet to be considered. The first of these concerns the potential interactive effects familiarity with an observed person may exert over the measurement of sexual interest. When considering explicit self-report findings, evidence suggests that more familiar people are considered to be more attractive across a range of methods, such as mere exposure paradigms (Han et al., 2020; Moreland & Beach, 1992; Rhodes et al., 2005). This suggests that familiarity can influence judgements of attractiveness at an *explicit* level. However, the extent to which this may occur when considering *implicit* measurement remained largely unexplored. Some evidence has highlighted the potential ability of these measures for assessing familiarity, and much of this work has focused primarily on pupillary responses

(Kafkas & Montaldi, 2015; Naber et al., 2013; Weiss et al., 2016). Specifically, the pupil old/ new effect has been identified, whereby pupil size is larger when familiar stimuli are viewed in comparison with novel stimuli (Võ et al., 2008). However, familiarity has not been explored alongside sexual interest using such implicit paradigms. As such, this thesis aimed to explore how these two factors may interact using a range of implicit cognitive measures.

The second aspect of measures of sexual interest that is investigated in this thesis is the sensitivity of such responses in distinguishing between attraction at a *within* sex category level, in comparison to the *between* category level. Until recently, the majority of implicit empirical investigation has focused on observing sexual interest in relation to sexual orientation, by contrasting differences in responses during the viewing of male and female images. This is evident when considering a range of implicit measures, such as pupillary responses (Attard-Johnson et al., 2021; Rieger & Savin-Williams, 2012; Snowden et al., 2019) and Stroop tasks (Ciardha & Gormley, 2012; Ó Ciardha & Gormley, 2013). In paradigms where within-category differences have been investigated, such as viewing time tasks (Ebsworth & Lalumière, 2012; Gress, 2005; Rönspies et al., 2015), results have been contradictory (Israel & Strassberg, 2009; Lippa, 2012; Xu et al., 2017). Therefore, this thesis aimed to clarify these findings by investigating further whether implicit cognitive measures are sensitive to differences in attractiveness *within* as well as *between* sex categories.

The final aim of this thesis was to address the effectiveness of implicit measures of sexual interest while under pressure from manipulation attempts. Specifically, as pupillary responses are held to reflect activity within the autonomic nervous system (McDougal & Gamlin, 2015; Zele & Gamlin, 2020), the question arises of whether these responses are impervious to conscious control. While the manipulability of this response has not been widely examined, some evidence suggests that internal intentional control of attention can exert influence over pupil size when considered separately from sexual interest (Brink et al., 2016; Ehlers et al., 2016; Grandchamp et al., 2014). Furthermore, external forms of manipulation, such as differences in experimenter temperament, may also influence such responses (Chapman et al., 1969). However, contradictory results have also been observed, whereby pupil size changes to familiar stimuli still occur during feigned amnesia (Heaver & Hutton, 2011). Despite these findings, the manipulability of the pupil response has yet to be explored concurrently with sexual interest. Therefore, this thesis also aimed to investigate this to further understanding of pupil size as a robust measure of sexual interest.

5.1 Explicit Self-Report Measures

Chapter 2 explored the potential interactions between sexual interest and familiarity, alongside the secondary factor of liking, which refers to the affinity we hold for other people. Previous research has identified that these factors can interact across a range of measures (Brockner & Swap, 1976; Joseph, 1982; Moreland & Beach, 1992). This chapter aimed to further explore these factors through a self-report questionnaire, providing foundational knowledge for how familiarity, attractiveness and liking interact within explicit tasks. These findings were then built upon in consequential research examining these factors utilising implicit cognitive measures.

The results from Experiment 1 indicated that familiarity, attractiveness and liking are interrelated concepts, with positive correlations found between these factors when rating stimuli consisting of celebrity *names*. However, these interactions did not extend to ratings of celebrity *faces*. Familiarity, attractiveness and liking were related when considering the faces of people who were familiar to the observer. However, no relationship was observed for unfamiliar faces. As such, this suggests that judgements of attractiveness and liking are dependent on observers' familiarity with an individual when viewing facial images.

This disparity between results found for name and face stimuli suggests that familiarity, liking and attractiveness act as *dependent* factors when evaluating names and *independent* factors when viewing facial images. The divergence in the way face and name stimuli are treated might be because of the different familiarity cues that are provided by each stimulus type. For example, names provide no direct visual information to facilitate the recognition of a familiar person. Thus, any judgments of a person's appearance from their name must draw on internal cognitive representations of an identity. In contrast, facial images provide more cues as observers may recognise someone without knowing them acutely enough to recall a name. As such, the differences in the information provided by each stimulus type may drive these effects, whereby familiarity interacts with liking and attractiveness when considering name stimuli and familiar, but not unfamiliar, facial images. In order to fully understand how these stimulus types may differ, further research aiming to compare name and face stimuli directly, as well as in isolation, would be beneficial and help develop the findings of this experiment.

It may be that the expected interactions between attractiveness and liking are only present within familiar others. When viewing an unfamiliar name, it is impossible to garner feelings of attractiveness and liking, as there is no information present on which to base these judgements. Yet, this is not the case for facial stimuli. As ratings of attractiveness and liking can be made based on the image provided.

Therefore, this experiment provides evidence suggesting that factors of familiarity, attractiveness and liking are only correlated when considering stimuli familiar to the observer.

An unexpected secondary pattern of results emerged from this experiment. Within the attractiveness factor, the previously described relationships only occurred when considering how personally attractive a person is, rather than how generally attractive they are. This trend of results was identified for both name and facial stimuli. This suggests that explicit measures of attractiveness are only sensitive to potential interactions with secondary factors, such as familiarity and liking, when considering personal attraction. As such, this furthers the theoretical understanding of what the relationships between factors of attractiveness, liking and familiarity are reflecting, namely that these factors are inter-related when considering these concepts as they personally relate to an individual, and not at a more objective, general level. This finding informed later chapters in this thesis, with all further experiments measuring attractiveness judgements relative to the participant, rather than as a general concept, in order to effectively assess the relationships between this factor and other factors of interest, such as familiarity.

5.2 Implicit Measures of Sexual Interest

Chapter 2 provided a basis of understanding for how the concepts of familiarity and attractiveness may interact. Chapter 3 aimed to expand on these findings and explore these factors using implicit cognitive measures. A range of paradigms were employed, such as viewing time (Experiment 2, 3 and 4) and Stroop tasks (Experiment 5). Differences in attraction reflective of sexual interest have been identified throughout the wider literature (Ciardha & Gormley, 2012; Gress, 2005;

Rönspies et al., 2015). This finding was corroborated in Experiment 3, when participants were presented with full body images. However, Experiment 2 did not identify this trend when employing facial stimuli. This suggests that facial stimuli, as opposed to full body images, may not have been effective in producing differences in attractiveness between sex categories within this chapter. This may be due to the lack of sexually salient content present in the face in comparison to the rest of the body. While full body stimuli contain sexually interesting features, such as the genital and breast areas, facial stimuli do not consist of such sexually salient imagery. As such, it may be that the lack of sexual features within facial stimuli restrict any potential attractiveness differences between sex categories.

Alongside investigating *between* sex category attractiveness, Chapter 3 aimed to explore any potential differences in attraction *within* these categories. There exists limited exploration into whether implicit cognitive measures are sensitive to the fine-grained differences which occur *within* individuals of the same sex, with such experiments reporting contradictory results (Israel & Strassberg, 2009; Lippa, 2012; Lippa et al., 2010). Within sex category differences in attractiveness were directly investigated in Experiments 3, 4 and 5 through within-person variability, whereby each identity was represented through two images, one in swimwear and one in casual clothing. Across experiments, no within sex category effects were identified through implicit cognitive measurement. However, the explicit ratings of attractiveness provided alongside each cognitive task showed a consistent pattern of results, with people in swimwear generally considered to be more attractive that their casually clothed counterparts. Due to the rating data indicating within sex category differences, it can be assumed that the stimuli did vary in attractiveness within identities. As such, the absence of such differences within the implicit cognitive measures suggests that these measures are not sensitive to differences in attractiveness within sex categories, a finding further corroborated due to the consistency of results across multiple stimulus sets.

It should be noted that the discussed pattern of results was identified only within male participants. Experiments 2 and 3 included both a male and a female observer population, however the female observers showed conflicting results within both explicit and implicit measurement, which occurred at both the *between* and *within* sex category level. This pattern of complicated female response aligns with those found across the wider literature (Bailey, 2009; Chivers et al., 2004), whereby female sexual response is incongruent with sexual orientation.

Taken together, the results from Chapter 3 indicate that implicit cognitive measures of sexual interest, as assessed through attractiveness, are reflective of sexual orientation within male observers. This finding was specific for full body stimuli, but not evident when viewing only facial images. While these responses were indicative of sexual interest *between* sex categories, no differences were identified *within* such categories. Therefore, this suggests that implicit cognitive measures of sexual interest are not sensitive to the more nuanced differences present between individuals of the same sex, despite reflecting sexual interest between those of different sexes. However, it should be noted that post-hoc analysis indicated that the statistical power of these experiments were limited. Therefore, to increase confidence in the findings from this chapter, it would be pertinent to re-run these experiments with an increased participant sample. Furthermore, the attractiveness ratings provided for the casual and beachwear images throughout these experiments were relatively similar, while not a ceiling. As such, it is possible that this limited the potential role of attractiveness on viewing times. A direct replication utilising stimuli which differ more drastically in average attractiveness ratings would be advantageous to corroborate the findings from these experiments.

5.3 Manipulability of Pupillary Responses of Sexual Interest

The final empirical chapter of this thesis, Chapter 4, aimed to investigate to what extent implicit cognitive measures are manipulatable. Specifically, the autonomic measure of pupil size was investigated to identify if the widely observed between-sex category differences in attraction for male participants (Attard-Johnson et al., 2021; Hess et al., 1965; Rieger & Savin-Williams, 2012) can be intentionally controlled. Experiment 6 aimed to explore how participant led manipulation of attention can affect pupillary responses of sexual interest. In line with the suggestion that differences in attention can influence pupil size (Brink et al., 2016; Ehlers et al., 2016; Grandchamp et al., 2014), participants were asked to focus their attention on specific areas of the display, comprising either the person, the background or a box on the background. Fixation locations indicated that observers generally adhered to instructions across conditions. Furthermore, despite the presence of a person image, participants were able to resist attending to these targets when instructed to direct attention elsewhere.

When considering pupillary responses, participants' pupil size was comparable when viewing same and opposite sex stimuli, regardless of which region on the image was viewed. The similar pupillary responses across sex category initially suggest that implicit measures of sexual interest can be manipulated based on redirection of attention. However, these effects were also absent when participants directly viewed the person stimuli. Therefore, it is possible that responses reflective of sexual interest were masked by task demands. For example, by intentionally directing attention, a task intrinsically linked with working memory load (Kiyonaga & Egner, 2013; Oberauer, 2019; Soto et al., 2005), this primary task demand may have limited the capacity of working memory available for judgements of sexual interest. As evidence shows that trait inferences require working memory (Wells et al., 2011; Zhou et al., 2021), this may have led to diminished responses reflective of such interest. As such, it seems plausible that experimental designs requiring participants to undertake an attentional task may diminish responses indicative of sexual interest in comparison to tasks where participants are able to view stimuli freely, a paradigm through which pupillary responses of sexual interest have been consistently identified (Attard-Johnson et al., 2016, 2017; Rieger et al., 2015).

While Experiment 6 investigated the potential effects of participant-led attention-based control on pupillary responses of sexual interest, Experiment 7 pivoted to explore other possible forms of manipulation, namely the presence of an examiner. One study has explored the potential effects of experimenter personality on pupil responses of sexual interest (Chapman et al., 1969) and identified that such responses were diminished by the presence of a formal, business-like examiner and were facilitated by a more casual individual. This opened the question of how pupil size would respond to sexual interest in the presence of an experimenter *per se*, which Experiment 7 explored by manipulating whether an experimenter was present or absent while viewing magazine covers featuring partially clothed individuals.

Attractiveness ratings were gathered in this experiment and indicated that sexual interest differences were present at an explicit self-report level, whereby ratings were congruent with sexual orientation. When considering viewing patterns, the male participants viewed images of females more than males, and the background of stimuli containing males more than females. These differences in viewing pattern were consistent regardless of whether an experimenter was present or absent. Despite this, larger pupil size was elicited when the experimenter was present in the room. This may be due to a period of increased attention occurring when observers were aware that they were being watched. This suggestion is supported by previous evidence whereby larger pupil dilation is found in relation to attentional effort (Alnæs et al., 2014; Kang et al., 2014; Karatekin et al., 2004). It is also possible that the presence of an experimenter led to increased arousal, subsequently leading to an increase in pupil size, a phenomenon also indicated through prior literature (Aboyoun & Dabbs, 1998; Snowden et al., 2019), and explored further later in this chapter.

When considering sex categories, larger pupil size were found for male than female stimuli in this experiment, a result inconsistent with the wider literature (Attard-Johnson et al., 2021; Ó Ciardha et al., 2018; Rieger et al., 2015). However, general pupil size was large throughout this experiment, especially when considering those identified in Experiment 6. As such, it may be that pupillary responses of sexual interest were influenced by a secondary factor which increased pupil size across all stimuli. There is some evidence which suggests that sexual arousal can lead to larger pupil dilation regardless of target sex (Aboyoun & Dabbs, 1998; Snowden et al., 2019). Therefore, it is possible that responses pertaining to sexual interest in Experiment 7 were masked due to the general influence of arousal. Taken together, Experiment 6 and 7 suggest that pupillary responses of sexual interest may be affected by factors such as attentional demands and general arousal. As such, future experiments aiming to explore attraction through this measure must consider that pupil size is not impervious to potential confounds. However, akin to the pervious chapter, the experiments contained in Chapter 4 had low statistical power, as indicated by post-hoc power analysis. As such, in order to add credence to the suggestions indicated by the results within this chapter, a larger sample size will be required.

5.4 Theoretical Implications

When considering the empirical findings from Chapters 2, 3 and 4 of this thesis, three distinct theoretical contributions arise. The first pertains to the idea that cognitive measures of sexual interest are actually evaluating a general assessment of mate preference rather than judgements of attractiveness. Secondly, the results from this thesis suggest that implicit cognitive measures of sexual interest may not be as effective as previously thought, as such responses can be potentially eliminated by secondary factors, such as attentional demands and high general arousal. Finally, it may be that cognitive measures of sexual interest are simply not as strong as initially believed, due to the inconsistent pattern of between-sex category results found across the experiments in this thesis.

5.4.1 MATE PREFERENCE

Chapter 3 explored how target attractiveness interacts with implicit cognitive measurements at a *between* and a *within* sex level. The only significant effects identified were for whole body stimuli when considering attractiveness differences between sex, whereby male participants viewed female targets for longer than male targets. Despite the explicit self-report data indicating that models in swimwear were considered more attractive than those in casualwear, this was not mirrored through implicit cognitive measurement. As such, it is plausible that these paradigms are not assessing attractiveness, and rather are reflecting a broader category of attraction, such as mate preference. This term relates to the sexual interest present for a group of others, a simplistic response unable to distinguish between individuals within these groups.

Support for this theory is evident across the literature, whereby sexual interest is measured almost exclusively in relation to sexual orientation (Ebsworth & Lalumière, 2012; Rieger et al., 2015; Rönspies et al., 2015) or age preference (Attard-Johnson et al., 2017; Harris et al., 1996; Mokros et al., 2010). Both of these responses are category specific, in that they orient observers towards categories of people, supporting the idea that these measures are reflecting broad evaluations of mate preference. However, there has been some suggestion that implicit cognitive measures of sexual interest are able to measure attraction more specifically than at this broad level (Lippa, 2012; Lippa et al., 2010). In these experiments, longer viewing times were observed for more attractive stimuli models within sex categories. Conflicting findings have also been discovered, whereby *within* sex category effects were absent when *between* effects were present (Israel & Strassberg, 2009). While the methodologies employed in these papers were generally comparable, it is possible that these discrepancies can be accounted for by terminology differences. As discussed in Chapter 3, a range of terminologies have been utilised across the literature to explore attractiveness, such as sexual attractiveness (Harris et al., 1996; Quinsey et al., 1996; Worling, 2006), (general) attractiveness (Hassebrauck, 1998) and sexual appeal (Dawson et al., 2012; Ebsworth & Lalumière, 2012; Lalumière et al., 2018). While implicit cognitive measures were found to reflect attraction within sex categories when observers were asked to rate images based on how sexually attractive they found them (Lippa, 2012; Lippa et al., 2010), this finding was not identified when participants rated stimuli based on sexual appeal. Therefore, it remains plausible that the utilisation of different terminologies may have led to the measurement of separate factors, sexual attractiveness and sexual appeal, rather than one universal factor of attractiveness. Due to the inconsistencies in measurement terminology, it remains unknown whether implicit cognitive measures of attractiveness are sensitive to attraction within sex categories.

The findings from this thesis, specifically pertaining to Chapter 3, support the idea that cognitive measures of sexual interest are specific to attractiveness *between* sex categories and are not sensitive to the more fine-grained differences in attractiveness present *within* these categories. These findings suggest that what is actually being measured is a category specific mate preference, thus orientating us *towards* but not *between* groups of people. Therefore, this thesis informs theoretical understanding by exploring what exactly these paradigms are measuring, with suggestions that these measures are reflective of a general mate preference response and are not sensitive to more specific forms of attraction. As such, this thesis suggests a change in how we should think about implicit cognitive measures of sexual interest, whereby such responses are capable of tapping into simplistic determinations of interest *between* sex categories but not *within* categories of sexually attractive others.

5.4.2 SUSCEPTIBILITY TO MANIPULATION ATTEMPTS

Chapter 4 of this thesis aimed to explore how implicit pupillary responses of sexual interest may be influenced by manipulation attempts. This constitutes the first empirical exploration of how participant-driven control attempts, namely attention redirection, can influence pupil dilation reflective of sexual interest. When sexual interest is not considered, pupil size has been found to increase during periods of attentional effort (Alnæs et al., 2014; Kang et al., 2014; Karatekin et al., 2004). It was unknown whether these effects would extend to the measurement of a secondary factor, such as sexual interest. The results from Experiment 6 indicated that intentional control of attention can interfere with any potential responses driven by sexual interest. This suggests that the attention-based pupil size responses previously identified can extend to the measurement of secondary factors. Therefore, experiments utilising pupillary measures should be cognizant of the potentially confounding role of attention redirection. For example, these effects may be reduced by employing a methodology with limited viewing instructions, such as a freeviewing design whereby observers are able to view images however feels natural to them. This design has been implemented across a number of pupillary response experiments where sexual interest has been identified (Attard-Johnson et al., 2016, 2017; Rieger et al., 2015), and thus may be effective in limiting the confounding role of attentional effects on pupil size measures of sexual interest.

Furthermore, this chapter explored how external factors may influence this sexual interest response. Prior evidence suggested that experimenter personality can affect such responses through social desirability effects (Chapman et al., 1969). It was unknown how the removal of such effects, through the absence of an experimenter, could influence pupillary responses of sexual interest. Experiment 7 showed that pupil size was generally larger when an experimenter was present than absent. This may be due to an increase in attention effort, as participants were aware that their responses were being monitored when the experimenter was present in the room. However, within this experiment there was no relationship between experimenter presence and sex category responses, and no pupil size differences attributed to sex category in isolation. Further research will therefore be required to assess the role of experimenter presence on pupil size using a methodology previously found to be effective in producing sexual interest effects, such as utilising a stimulus set previously identified to be effective in producing such effects within pupil size measures. This would lend support to the findings outlined by this experiment.

Despite the reduction of sex category effects on pupillary responses of sexual interest found in Experiment 7, this itself provides clues surrounding the potential control of such responses. Average pupil size in Experiment 7 was unexpectedly large for all images regardless of target sex. This could be accounted for by individual differences in pupil size, whereby participants in Experiment 7 had generally large pupils regardless of presented stimuli or experimental design. However, as these participants, with two exceptions, also partook in Experiment 6 when average pupil size was much lower, this explanation is unlikely. Consequently, it is more likely that the heightened pupillary responses are reflective of a secondary factor, such as arousal. The potential role of arousal on pupil size has seen conflicting results within the measurement of sexual interest. Some evidence has found that high levels of sexual arousal produce comparably elevated pupil size regardless of target sex, with similar responses reported towards both male and female images (Aboyoun & Dabbs, 1998; Snowden et al., 2019). However, increased pupil size has been found in response to sexual interest in conditions of high arousal, whereby pupillary responses reflect sexual orientation (Rieger et al., 2015), and thus the role of arousal on pupil size change is unclear. It is possible that these differences may have arisen due to methodological differences, such as variations in stimuli choice. While pupillary responses reflective of sexual orientation have been identified during presentation of sexually arousing videos (Rieger et al., 2015), in some instances this effect has not been found during the viewing of still images (Snowden et al., 2019). The extent to which these differences in effects may be stimulus-driven remains uncertain as some experiments have successfully identified sexual orientation while employing still images (Attard-Johnson et al., 2017; Attard-Johnson & Bindemann, 2017; Hess et al., 1965). As such, to address these inconsistent findings, further experiments aiming to investigate the effects of sexual arousal over pupillary responses of sexual interest are necessary. For example, an experiment which systematically varies stimuli arousal level, across stimuli type, may help to address this gap in knowledge.

When considering this thesis, the findings support the theory that arousal leads to general increases in pupil size indiscriminate of target sex. Therefore, while arousal was not explicitly investigated throughout this thesis, results suggest that this factor may exert influence over pupillary responses of sexual interest. Overall, this thesis suggests that pupil size responses indicative of sexual interest are potentially susceptible to intentional manipulation. These can range from observer led forms of control, such as attention redirection, to methodological design elements, such as experimenter effects and stimulus driven arousal. As such, care should be taken when considering pupillary responses of sexual interest to ensure that results are reflective of the factor of interest and not a potential confound, due to their apparent susceptibility to manipulation.

5.4.3 RELIABILITY OF IMPLICIT MEASUREMENT OF SEXUAL INTEREST

As aforementioned, implicit cognitive measures of sexual interest may not be as reliable as previously thought due to potential susceptibility to manipulation attempts. This raises the question of how strong these effects are as a whole. Throughout this thesis, the expected effects reflective of sexual orientation were identified in only one experiment, Experiment 3. However, sex category effects were explored in a total of four experiments throughout this thesis. This was despite the implementation of various stimulus sets and methodological design. As such, this points to the suggestion that these effects may not be as reliably present as evidence suggests.

There are two main explanations for the inconsistent sexual interest results identified throughout this thesis. The first being that secondary factors may have confounded or diminished potential response, a possibility explored in the previous section. A second explanation may be publication bias, which refers to the phenomenon occurring when published studies have systematically different findings to unpublished experiments (Song et al., 2013). While these effects are hard to distinguish, meta-analyses can aid in uncovering potential bias. When considering implicit cognitive measures of sexual interest, publication bias has been explored in relation to pupillary responses via a meta-analysis (Attard-Johnson et al., 2021). This study explored publication bias by investigating the possibility of selective reporting practices, such as p-hacking, and by exploring the relationship between a study's precision and effect size. While some evidence of bias was identified when considering male sexual response, it was identified that sex category effects persisted, albeit to a lesser extent, when accounting for such bias. While it remains unknown the extent to which publication bias may be present within other forms of implicit cognitive measurement, this finding suggests that sexual interest findings are likely not accounted for in their entirety by publication bias. Therefore, it seems evident that the absence of sexual orientation congruent responses throughout this thesis may be accounted for by the influence of secondary factors or due to experimental design.

5.5 Limitations and Future Directions

One limitation evident throughout this thesis was that experimental design did not allow for between sex category effects to be analysed for every experiment. Some of the experiments, (2, 3 and 4) included presentation of exclusive female targets to heterosexual male observers. While this allowed for the investigation of potential within-sex category differences in attractiveness, it meant that it was not possible to ensure that participants showed responses reflective of sexual interest at the between-sex category level. The measurement of these responses may have lent strength to the conclusions drawn from the results. Future research should focus on exploring both *between* and *within* sex category differences in attractiveness in tandem to further theoretical understanding pertaining to the sensitivity of implicit cognitive measurement to such effects.

Secondly, the caveat of arousal present in Experiment 7, whereby pupil size was elevated across all stimuli, was not accounted for through experimental design. While sexual interest was explored through implicit and explicit tasks, there was no measurement of general arousal collected. This meant that further exploration of the role of this factor was not possible and as such it remains unknown whether arousal masks sex category effects or eliminates these responses completely. Further research should be undertaken to explore the potential interactions of arousal and sexual interest on pupil size to investigate the nature of these results. For example, we could gain a better understanding of how pupillary responses may interact with arousal by presenting observers with stimuli varying in sexually and generally arousing content, such as clothed images and pornography, and measuring perceived levels of arousal through pupil size and rating tasks. Furthermore, it would be interesting to see whether any interactions carry over to other forms of cognitive measurement, such as the viewing time and Stoop tasks explored throughout this thesis.

Finally, both pupillary response experiments within Chapter 4 would benefit from further investigation. As discussed, the arousal caveat identified in Experiment 7 should be addressed in a follow-up experiment. Experiment 6 identified that intentional attentional redirection can potentially diminish pupil size in relation to sexual interest. However, as the expected between sex category effects were not identified when participants were able to freely view the person target, it remains unknown as to whether the mere presence of general attentional task demands can account for this effect. A direct replication of this experiment would be useful in examining whether these effects are consistent. Furthermore, studies aiming to vary the strength of attention-based demands, for example, by systematically varying attentional effort through increases in instructional demands, would aid in furthering the theoretical understanding of the manipulability of pupillary responses of sexual interest.

References

Abel, G. G., Huffman, J., Warberg, B., & Holland, C. L. (1998). Visual Reaction Time and Plethysmography as Measures of Sexual Interest in Child Molesters. *Sexual Abuse: A Journal of Research and Treatment*, *10*(2), 81–95. https://doi.org/10.1023/A:1022063214826

Abel, G. G., Jordan, A., Rouleau, J. L., Emerick, R., Barboza-Whitehead, S., & Osborn, C. (2004). Use of Visual Reaction Time to Assess Male Adolescents Who Molest Children. *Sexual Abuse: A Journal of Research and Treatment*, *16*(3), 255–265. https://doi.org/10.1023/B:SEBU.0000029136.01177.51

Aboyoun, D. C., & Dabbs, J. M. (1998). THE HESS PUPIL DILATION FINDINGS: SEX OR NOVELTY? *Social Behavior and Personality: An International Journal*, *26*(4), 415–419. https://doi.org/10.2224/sbp.1998.26.4.415

Akerman, G., & Beech, A. R. (2012). A Systematic Review of Measures of Deviant Sexual Interest and Arousal. *Psychiatry, Psychology and Law*, *19*(1), 118–143. https://doi.org/10.1080/13218719.2010.547161

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4), 1. https://doi.org/10.1167/14.4.1

Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. M. (2015). Telling faces together: Learning new faces through exposure to multiple instances. *Quarterly Journal of Experimental Psychology*, *68*(10), 2041–2050. https://doi.org/10.1080/17470218.2014.1003949

Ashley, V., & Swick, D. (2009). Consequences of emotional stimuli: Age differences on pure and mixed blocks of the emotional Stroop. *Behavioral and Brain Functions*, *5*(1), 14. https://doi.org/10.1186/1744-9081-5-14

Ashwin, C., Wheelwright, S., & Baron-Cohen, S. (2006). Attention bias to faces in Asperger Syndrome: A pictorial emotion Stroop study. *Psychological Medicine*, *36*(6), 835–843. https://doi.org/10.1017/S0033291706007203

Attard-Johnson, J., & Bindemann, M. (2017). Sex-specific but not sexually explicit: Pupillary responses to dressed and naked adults. *Royal Society Open Science*, *4*(5), 160963. https://doi.org/10.1098/rsos.160963

Attard-Johnson, J., Bindemann, M., & Ó Ciardha, C. (2016). Pupillary Response as an Age-Specific Measure of Sexual Interest. *Archives of Sexual Behavior*, *45*(4), 855–870. https://doi.org/10.1007/s10508-015-0681-3

Attard-Johnson, J., Bindemann, M., & Ó Ciardha, C. (2017). Heterosexual, Homosexual, and Bisexual Men's Pupillary Responses to Persons at Different Stages of Sexual Development. *The Journal of Sex Research*, *54*(9), 1085–1096. https://doi.org/10.1080/00224499.2016.1241857

Attard-Johnson, J., Ó Ciardha, C., & Bindemann, M. (2019). Comparing methods for the analysis of pupillary response. *Behavior Research Methods*, *51*(1), 83–95. https://doi.org/10.3758/s13428-018-1108-6

Attard-Johnson, J., Vasilev, M. R., Ó Ciardha, C., Bindemann, M., & Babchishin, K. M. (2021). Measurement of Sexual Interests with Pupillary Responses: A Meta-Analysis. *Archives of Sexual Behavior*, *50*(8), 3385–3411. https://doi.org/10.1007/s10508-021-02137-y

Atwood, R. W., & Howell, R. J. (1971). Pupillometric and personality test score differences of female aggressing pedophiliacs and normals. *Psychonomic Science*, *22*(2), 115–116. https://doi.org/10.3758/BF03332522

Bailey, J. M. (2009). What is Sexual Orientation and Do Women Have One? In D. A. Hope (Ed.), *Contemporary Perspectives on Lesbian, Gay, and Bisexual Identities* (pp. 43–63). Springer. https://doi.org/10.1007/978-0-387-09556-1_3

Baker, K. A., Laurence, S., & Mondloch, C. J. (2017). How does a newly encountered face become familiar? The effect of within-person variability on adults' and children's perception of identity. *Cognition*, *161*, 19–30. https://doi.org/10.1016/j.cognition.2016.12.012

Barbaree, H. E., & Marshall, W. L. (1988). Deviant sexual arousal, offense history, and demographic variables as predictors of reoffense among child molesters. *Behavioral Sciences & the Law*, *6*(2), 267–280. https://doi.org/10.1002/bsl.2370060209

Barbaree, H. E., Marshall, W. L., & Lanthier, R. D. (1979). Deviant sexual arousal in rapists. *Behaviour Research and Therapy*, *17*(3), 215–222. https://doi.org/10.1016/0005-7967(79)90036-6

Barker, J. G., & Howell, R. J. (1992). The Plethysmograph: A Review of Recent Literature. *Bull Am Acad Psychiatry Law*, *20*(1).

Becker, E. S., Rinck, M., Margraf, J., & Roth, W. T. (2001). The emotional Stroop effect in anxiety disorders: General emotionality or disorder specificity? *Journal of Anxiety Disorders*, *15*(3), 147–159. https://doi.org/10.1016/S0887-6185(01)00055-X

Bernick, N., Kling, A., & Borowitz, G. (1971). Physiologic Differentiation of Sexual Arousal and Anxiety. *Psychosomatic Medicine*, *33*(4), 341.

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision*, *9*(2), 7. https://doi.org/10.1167/9.2.7

Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, A. M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1477–1485. https://doi.org/10.1037/a0019057

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

Brink, R. L. van den, Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil Diameter Tracks Lapses of Attention. *PLOS ONE*, *11*(10), e0165274. https://doi.org/10.1371/journal.pone.0165274

Brockner, J., & Swap, W. C. (1976). Effects of repeated exposure and attitudinal similarity on self-disclosure and interpersonal attraction. *Journal of Personality and Social Psychology*, *33*(5), 531–540. https://doi.org/10.1037/0022-3514.33.5.531

Brown, M., Amoroso, D. M., Ware, E. E., Pruesse, M., & Pilkey, D. W. (1973). Factors Affecting Viewing Time of Pornography. *The Journal of Social Psychology*, *90*(1), 125–135. https://doi.org/10.1080/00224545.1973.9922628

Burley, D. T., Gray, N. S., & Snowden, R. J. (2017). As Far as the Eye Can See: Relationship between Psychopathic Traits and Pupil Response to Affective

Stimuli. *PLOS ONE*, *12*(1), e0167436. https://doi.org/10.1371/journal.pone.0167436

Burley, D. T., Gray, N. S., & Snowden, R. J. (2019). Emotional modulation of the pupil response in psychopathy. *Personality Disorders: Theory, Research, and Treatment*, *10*(4), 365–375. https://doi.org/10.1037/per0000313

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology (London, England: 1953)*, *81 (Pt 3)*, 361–380. https://doi.org/10.1111/j.2044-8295.1990.tb02367.x

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity From Variation: Representations of Faces Derived From Multiple Instances. *Cognitive Science*, *40*(1), 202–223. https://doi.org/10.1111/cogs.12231

Burton, M., A. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66*(8), 1467–1485. https://doi.org/10.1080/17470218.2013.800125

Chapman, L. J., Chapman, J. P., & Brelje, T. (1969). Influence of the experimenter on pupillary dilation to sexually provocative pictures. *Journal of Abnormal Psychology*, *74*(3), 396–400. https://doi.org/10.1037/h0027493

Chen, H., Dey, A., Mark, B., & Lindeman, R., W. (2017). *Exploring Pupil Dilation in Emotional Virtual Reality Environments*. https://ir.canterbury.ac.nz/handle/10092/15404

Chetverikov, A., & Upravitelev, P. (2016). Online versus offline: The Web as a medium for response time data collection. *Behaviour research methods, 48,* 1086-1099. https://doi.org/10.3758/s13428-015-0632-x

Chivers, M. L., & Bailey, J. M. (2005). A sex difference in features that elicit genital response. *Biological Psychology*, *70*(2), 115–120. https://doi.org/10.1016/j.biopsycho.2004.12.002

Chivers, M. L., Rieger, G., Latty, E., & Bailey, J. M. (2004). A Sex Difference in the Specificity of Sexual Arousal. *Psychological Science*, *15*(11), 736–744. https://doi.org/10.1111/j.0956-7976.2004.00750.x

Chivers, M. L., Seto, M. C., & Blanchard, R. (2007). Gender and sexual orientation differences in sexual response to sexual activities versus gender of actors in sexual films. *Journal of Personality and Social Psychology*, *93*(6), 1108–1121. https://doi.org/10.1037/0022-3514.93.6.1108

Chivers, M. L., Seto, M. C., Lalumière, M. L., Laan, E., & Grimbos, T. (2010). Agreement of Self-Reported and Genital Measures of Sexual Arousal in Men and Women: A Meta-Analysis. *Archives of Sexual Behavior*, *39*(1), 5–56. https://doi.org/10.1007/s10508-009-9556-9

Ciardha, C. Ó., & Gormley, M. (2009). Comparing two implicit cognitive measures of sexual interest: A pictorial modified Stroop task and the implicit association test. In *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders* (pp. 177–201). John Wiley & Sons. https://doi.org/10.1002/9780470747551

Ciardha, C. Ó., & Gormley, M. (2012). Using a Pictorial-Modified Stroop Task to Explore the Sexual Interests of Sexual Offenders Against Children. *Sexual Abuse*, *24*(2), 175–197. https://doi.org/10.1177/1079063211407079

Cisler, J. M., Wolitzky-Taylor, K. B., Adams, T. G., Babson, K. A., Badour, C. L., & Willems, J. L. (2011). The emotional Stroop task and posttraumatic stress disorder: A meta-analysis. *Clinical Psychology Review*, *31*(5), 817–828. https://doi.org/10.1016/j.cpr.2011.03.007

Clift, R. J. W., Rajlic, G., & Gretton, H. M. (2009). Discriminative and Predictive Validity of the Penile Plethysmograph in Adolescent Sex Offenders. *Sexual Abuse*, *21*(3), 335–362. https://doi.org/10.1177/1079063209338491

Commisso, M., & Finkelstein, L. (2012). Physical Attractiveness Bias in Employee Termination. *Journal of Applied Social Psychology*, *42*(12), 2968–2987. https://doi.org/10.1111/j.1559-1816.2012.00970.x

Constantine, R., McNally, R. J., & Hornig, C. D. (2001). Snake Fear and the Pictorial Emotional Stroop Paradigm. *Cognitive Therapy and Research*, *25*(6), 757–764. https://doi.org/10.1023/A:1012923507617

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioural research. *PloS one, 8*(3), e57410. https://doi.org/10.1371/journal.pone.0057410

Dawson, S. J., & Chivers, M. L. (2016). Gender-Specificity of Initial and Controlled Visual Attention to Sexual Stimuli in Androphilic Women and Gynephilic Men. *PLOS ONE*, *11*(4), e0152785. https://doi.org/10.1371/journal.pone.0152785

Dawson, S. J., Suschinsky, K. D., & Lalumière, M. L. (2012). Sexual Fantasies and Viewing Times Across the Menstrual Cycle: A Diary Study. *Archives of Sexual Behavior*, *41*(1), 173–183. https://doi.org/10.1007/s10508-012-9939-1

Dietz, J., Bradley, M. M., Okun, M. S., & Bowers, D. (2011). Emotion and ocular responses in Parkinson's disease. *Neuropsychologia*, *49*(12), 3247–3253. https://doi.org/10.1016/j.neuropsychologia.2011.07.029

Dombert, B., Mokros, A., Brückner, E., Schlegl, V., Antfolk, J., Bäckström, A., Zappalà, A., Osterheider, M., & Santtila, P. (2013). The Virtual People Set: Developing Computer-Generated Stimuli for the Assessment of Pedophilic Sexual Interest. *Sexual Abuse*, *25*(6), 557–582. https://doi.org/10.1177/1079063212469062

Dombert, B., Schmidt, A. F., Banse, R., Briken, P., Hoyer, J., Neutze, J., & Osterheider, M. (2016). How Common is Men's Self-Reported Sexual Interest in Prepubescent Children? *The Journal of Sex Research*, *53*(2), 214–223. https://doi.org/10.1080/00224499.2015.1020108

Dresler, T., Attar, C. H., Spitzer, C., Löwe, B., Deckert, J., Büchel, C., Ehlis, A.-C., & Fallgatter, A. J. (2012). Neural correlates of the emotional Stroop task in panic disorder patients: An event-related fMRI study. *Journal of Psychiatric Research*, *46*(12), 1627–1634. https://doi.org/10.1016/j.jpsychires.2012.09.004

Dresler, T., Ehlis, A.-C., Hindi Attar, C., Ernst, L. H., Tupak, S. V., Hahn, T., Warrings, B., Markulin, F., Spitzer, C., Löwe, B., Deckert, J., & Fallgatter, A. J. (2012). Reliability of the emotional Stroop task: An investigation of patients with panic disorder. *Journal of Psychiatric Research*, *46*(9), 1243–1248. https://doi.org/10.1016/j.jpsychires.2012.06.006

Dresler, T., Mériau, K., Heekeren, H. R., & van der Meer, E. (2009). Emotional Stroop task: Effect of word arousal and subject anxiety on emotional interference. *Psychological Research PRPF*, *73*(3), 364–371. https://doi.org/10.1007/s00426-008-0154-6

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*(1), 109–128. https://doi.org/10.1037/0033-2909.110.1.109

Ebsworth, M., & Lalumière, M. L. (2012). Viewing Time as a Measure of Bisexual Sexual Interest. *Archives of Sexual Behavior*, *41*(1), 161–172. https://doi.org/10.1007/s10508-012-9923-9

Efrain, M. G., & Patterson, E. W. J. (1974). Voters vote beautiful: The effect of physical appearance on a national election. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *6*(4), 352–356. https://doi.org/10.1037/h0081881

Ehlers, J., Bubalo, N., Loose, M., & Huckauf, A. (2015). *Towards Voluntary Pupil Control | Proceedings of the 2nd International Conference on Physiological Computing Systems*. https://dl.acm.org/doi/abs/10.5220/0005240000050012

Ehlers, J., Strauch, C., Georgi, J., & Huckauf, A. (2016). Pupil Size Changes as an Active Information Channel for Biofeedback Applications. *Applied Psychophysiology and Biofeedback*, *41*(3), 331–339. https://doi.org/10.1007/s10484-016-9335-z

Ekman, I., Poikola, A., Mäkäräinen, M., Takala, T., & Hämäläinen, P. (2008). Voluntary pupil size change as control in eyes only interaction. *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, 115–118. https://doi.org/10.1145/1344471.1344501

Ellis, C., J. (1981). The pupillary light reflex in normal subjects. | British Journal of Ophthalmology. *British Journal of Ophthalmology*, *65*(11), 754–759. http://dx.doi.org/10.1136/bjo.65.11.754

Epp, A. M., Dobson, K. S., Dozois, D. J. A., & Frewen, P. A. (2012). A systematic meta-analysis of the Stroop task in depression. *Clinical Psychology Review*, *32*(4), 316–328. https://doi.org/10.1016/j.cpr.2012.02.005

Erwin, P. G. (1993). First names and perceptions of physical attractiveness. *The Journal of Psychology: Interdisciplinary and Applied*, *127*(6), 625–631. https://doi.org/10.1080/00223980.1993.9914901

Finkel, E. J., Norton, M. I., Reis, H. T., Ariely, D., Caprariello, P. A., Eastwick, P. W., Frost, J. H., & Maniaci, M. R. (2015). When Does Familiarity Promote Versus Undermine Interpersonal Attraction? A Proposed Integrative Model From Erstwhile Adversaries. *Perspectives on Psychological Science*, *10*(1), 3–19. https://doi.org/10.1177/1745691614561682

Franzen, P. L., Buysse, D. J., Dahl, R. E., Thompson, W., & Siegle, G. J. (2009). Sleep deprivation alters pupillary reactivity to emotional stimuli in healthy young adults. *Biological Psychology*, *80*(3), 300–305. https://doi.org/10.1016/j.biopsycho.2008.10.010

Franzen, P. L., Siegle, G. J., & Buysse, D. J. (2008). Relationships between affect, vigilance, and sleepiness following sleep deprivation. *Journal of Sleep Research*, *17*(1), 34–41. https://doi.org/10.1111/j.1365-2869.2008.00635.x

Frieze, I. H., Olson, J. E., & Russell, J. (1991). Attractiveness and Income for Men and Women in Management1. *Journal of Applied Social Psychology*, *21*(13), 1039–1057. https://doi.org/10.1111/j.1559-1816.1991.tb00458.x

Fromberger, P., Meyer, S., Kempf, C., Jordan, K., & Müller, J. L. (2015). Virtual Viewing Time: The Relationship between Presence and Sexual Interest in Androphilic and Gynephilic Men. *PLOS ONE*, *10*(5), e0127156. https://doi.org/10.1371/journal.pone.0127156

Gallagher-Duffy, J., MacKay, S., Duffy, J., Sullivan-Thomas, M., & Peterson-Badali, M. (2009). The Pictorial Fire Stroop: A Measure of Processing Bias for Fire-Related Stimuli. *Journal of Abnormal Child Psychology*, *37*(8), 1165–1176. https://doi.org/10.1007/s10802-009-9339-8

Gavas, R., Chatterjee, D., & Sinha, A. (2017). Estimation of cognitive load based on the pupil size dilation. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1499–1504. https://doi.org/10.1109/SMC.2017.8122826

Goldman, W., & Lewis, P. (1977). Beautiful is good: Evidence that the physically attractive are more socially skillful. *Journal of Experimental Social Psychology*, *13*(2), 125–130. https://doi.org/10.1016/S0022-1031(77)80005-X

Grandchamp, R., Braboszcz, C., & Delorme, A. (2014). Oculometric variations during mind wandering. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00031

Gray, S. R., Abel, G. G., Jordan, A., Garby, T., Wiegel, M., & Harlow, N. (2015). Visual Reaction Time™ as a Predictor of Sexual Offense Recidivism. *Sexual Abuse*, *27*(2), 173–188. https://doi.org/10.1177/1079063213502680

Green, M. L., Kraus, S. K., & Green, R. G. (1979). Pupillary Responses to Pictures and Descriptions of Sex-Stereotyped Stimuli. *Perceptual and Motor Skills*, *49*(3), 759–764. https://doi.org/10.2466/pms.1979.49.3.759

Greene, H. H., & Rayner, K. (2001). Eye movements and familiarity effects in visual search. *Vision Research*, *41*(27), 3763–3773. https://doi.org/10.1016/S0042-6989(01)00154-7

Gress, C. L. Z. (2005). Viewing time measures and sexual interest: Another piece of the puzzle. *Journal of Sexual Aggression*, *11*(2), 117–125. https://doi.org/10.1080/13552600500063666

Gress, C. L. Z., Anderson, J. O., & Laws, D. R. (2013). Delays in attentional processing when viewing sexual imagery: The development and comparison of two measures. *Legal and Criminological Psychology*, *18*(1), 66–82. https://doi.org/10.1111/j.2044-8333.2011.02032.x

Guillon, M., Dumbleton, K., Theodoratos, P., Gobbe, M., Wooley, C. B., & Moody, K. (2016). The Effects of Age, Refractive Status, and Luminance on Pupil Size. *Optometry and Vision Science*, *93*(9), 1093–1100. https://doi.org/10.1097/OPX.0000000000000893

Halberstadt, J., Rhodes, G., & Catty, S. R. (2003). Subjective and objective familiarity as explanations for the attraction to average faces. In *Advances in Psychological Research* (Vol. 22, pp. 35–49). Nova Publishers.

Hamel, R. F. (1974). Female Subjective and Pupillary Reaction to Nude Male and Female Figures. *The Journal of Psychology*, *87*(2), 171–175. https://doi.org/10.1080/00223980.1974.9915687

Han, S., Liu, S., Gan, Y., Xu, Q., Xu, P., Luo, Y., & Zhang, L. (2020). Repeated exposure makes attractive faces more attractive: Neural responses in facial attractiveness judgement. *Neuropsychologia*, *139*. https://doi.org/10.1016/j.neuropsychologia.2020.107365

Hanslmayr, S., Pastötter, B., Bäuml, K.-H., Gruber, S., Wimber, M., & Klimesch, W. (2008). The Electrophysiological Dynamics of Interference during the Stroop Task. *Journal of Cognitive Neuroscience*, *20*(2), 215–225. https://doi.org/10.1162/jocn.2008.20020

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66*(2), 348–362. https://doi.org/10.1037/0022-006X.66.2.348

Harris, G. T., Rice, M. E., Quinsey, V. L., & Chaplin, T. C. (1996). Viewing time as a measure of sexual interest among child molesters and normal heterosexual men. *Behaviour Research and Therapy*, *34*(4), 389–394. https://doi.org/10.1016/0005-7967(95)00070-4

Harrison, A. A. (1969). Exposure and popularity. *Journal of Personality*, *37*(3), 359–377. https://doi.org/10.1111/j.1467-6494.1969.tb01750.x

Hassebrauck, M. (1998). The Visual Process Method: A New Method to Study Physical Attractiveness. *Evolution and Human Behavior*, *19*(2), 111–123. https://doi.org/10.1016/S1090-5138(98)00002-6

Haupt, S., Axmacher, N., Cohen, M. X., Elger, C. E., & Fell, J. (2009). Activation of the caudal anterior cingulate cortex due to task-related interference in an auditory Stroop paradigm. *Human Brain Mapping*, *30*(9), 3043–3056. https://doi.org/10.1002/hbm.20731

Hayward, G., Goodwin, G. M., & Harmer, C. J. (2004). The role of the anterior cingulate cortex in the counting Stroop task. *Experimental Brain Research*, *154*(3), 355–358. https://doi.org/10.1007/s00221-003-1665-4

Heaver, B., & Hutton, S. B. (2011). Keeping an eye on the truth? Pupil size changes associated with recognition memory. *Memory*, *19*(4), 398–405. https://doi.org/10.1080/09658211.2011.575788

Hensley, W. E., & Spencer, B. A. (1985). The effect of first names on perceptions of female attractiveness. *Sex Roles*, *12*(7), 723–729. https://doi.org/10.1007/BF00287866

Hess, E. H., & Polt, J. M. (1960). *Pupil Size as Related to Interest Value of Visual Stimuli*. https://doi.org/10.1126/science.132.3423.349

Hess, E. H., Seltzer, A. L., & Shlien, J. M. (1965). Pupil response of hetero- and homosexual males to pictures of men and women: A pilot study. *Journal of Abnormal Psychology*, *70*(3), 165–168. https://doi.org/10.1037/h0021978

Hoon, E. F., Hoon, P. W., & Wincze, J. P. (1976). An inventory for the measurement of female sexual arousability: The SAI. *Archives of Sexual Behavior*, *5*(4), 291–300. https://doi.org/10.1007/BF01542081

Hoss, R. A., & Langlois, J. H. (2003). Infants prefer attractive faces. In *The development of face processing in infancy and early childhood: Current perspectives* (pp. 27–38). Nova Publishers.

Howell, J. R., Reynolds, C. F., Thase, M. E., Frank, E., Jennings, J. R., Houck, P. R., Berman, S., Jacobs, E., & Kupfer, D. J. (1987). Assessment of sexual function, interest and activity in depressed men. *Journal of Affective Disorders*, *13*(1), 61–66. https://doi.org/10.1016/0165-0327(87)90074-7

Ignaschchenkova, A., Dicke, P. W., Haarmeier, T., & Their, P. (2004). Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nature Neuroscience, 7*(1), 56-64. https://doi.org/10.1038/nn1169

Imhoff, R., Schmidt, A. F., Nordsiek, U., Luzar, C., Young, A. W., & Banse, R. (2010). Viewing Time Effects Revisited: Prolonged Response Latencies for Sexually Attractive Targets Under Restricted Task Conditions. *Archives of Sexual Behavior*, *39*(6), 1275–1288. https://doi.org/10.1007/s10508-009-9595-2

Israel, E., & Strassberg, D. S. (2009). Viewing Time as an Objective Measure of Sexual Interest in Heterosexual Men and Women. *Archives of Sexual Behavior*, *38*(4), 551–558. https://doi.org/10.1007/s10508-007-9246-4

Janisse, M. P. (1973). Pupil size and affect: A critical review of the literature since 1960. *Canadian Psychologist / Psychologie Canadienne*, *14*(4), 311–329. https://doi.org/10.1037/h0082230

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition*, *121*(3), 313–323. https://doi.org/10.1016/j.cognition.2011.08.001

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences*, *103*(45), 17048–17052. https://doi.org/10.1073/pnas.0605678103

Jones, A. L., & Kramer, R. S. S. (2015). Facial Cosmetics Have Little Effect on Attractiveness Judgments Compared with Identity. *Perception*, *44*(1), 79–86. https://doi.org/10.1068/p7904

Joseph, W. B. (1982). The Credibility of Physically Attractive Communicators: A Review. *Journal of Advertising*, *11*(3), 15–24. https://doi.org/10.1080/00913367.1982.10672807

Joshi, S., & Gold, J. I. (2020). Pupil size as a window on neural substrates of cognition. *Trends in cognitive sciences, 24*(6), 466-480. https://doi.org/10.1016/j.tics.2020.03.005

Kafkas, A., & Montaldi, D. (2015). The pupillary response discriminates between subjective and objective familiarity and novelty. *Psychophysiology*, *52*(10), 1305–1316. https://doi.org/10.1111/psyp.12471

Kaiser, D., & Nyga, K. (2020). Tracking cortical representations of facial attractiveness using time-resolved representational similarity analysis. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-74009-9

Kalmus, E., & Beech, A. R. (2005). Forensic assessment of sexual interest: A review. *Aggression and Violent Behavior*, *10*(2), 193–217. https://doi.org/10.1016/j.avb.2003.12.002

Kang, O. E., Huffer, K. E., & Wheatley, T. P. (2014). Pupil Dilation Dynamics Track Attention to High-Level Information. *PLOS ONE*, *9*(8), e102463. https://doi.org/10.1371/journal.pone.0102463

Kappes, C., & Bermeitinger, C. (2016). The Emotional Stroop as an Emotion Regulation Task. *Experimental Aging Research*, *42*(2), 161–194. https://doi.org/10.1080/0361073X.2016.1132890

Karatekin, C., Couperus, J. W., & Marcus, D. J. (2004). Attention allocation in the dual-task paradigm as measured through behavioral and psychophysiological responses. *Psychophysiology*, *41*(2), 175–185. https://doi.org/10.1111/j.1469-8986.2004.00147.x

Khanna, M. M., Badura-Brack, A. S., McDermott, T. J., Embury, C. M., Wiesman, A. I., Shepherd, A., Ryan, T. J., Heinrichs-Graham, E., & Wilson, T. W. (2017). Veterans with post-traumatic stress disorder exhibit altered emotional processing and attentional control during an emotional Stroop task. *Psychological Medicine*, *47*(11), 2017–2027. https://doi.org/10.1017/S0033291717000460

Kim, M., Beversdorf, D. Q., & Heilman, K. M. (2000). Arousal response with aging: Pupillographic study. *Journal of the International Neuropsychological Society*, *6*(3), 348–350. https://doi.org/10.1017/S135561770000309X

Kinsey, A. C., Pomeroy, W. B., & Martin, C. E. (1948). *Sexual behavior in the human male* (1948-01530-000). Saunders.

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, *20*(2), 228–242. https://doi.org/10.3758/s13423-012-0359-y

Kniffin, K. M., & Wilson, D. S. (2004). The effect of nonphysical traits on the perception of physical attractiveness: Three naturalistic studies. *Evolution and Human Behavior*, *25*(2), 88–101. https://doi.org/10.1016/S1090-5138(04)00006-6

Kramer, R. S. S., Manesi, Z., Towler, A., Reynolds, M. G., & Burton, A. M. (2018). Familiarity and Within-Person Facial Variability: The Importance of the Internal and External Features. *Perception*, *47*(1), 3–15. https://doi.org/10.1177/0301006617725242

Kun, A. L., Palinko, O., & Razumenić, I. (2012). Exploring the effects of size and luminance of visual targets on the pupillary light reflex. *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 183–186. https://doi.org/10.1145/2390256.2390287

Laan, E., Everaerd, W., & Evers, A. (1995). Assessment of female sexual arousal: Response specificity and construct validity. *Psychophysiology*, *32*(5), 476–485. https://doi.org/10.1111/j.1469-8986.1995.tb02099.x

Laeng, B., & Sulutvedt, U. (2014). The Eye Pupil Adjusts to Imaginary Light. *Psychological Science*, *25*(1), 188–197. https://doi.org/10.1177/0956797613503556

Lalumière, M. L., Babchishin, K. M., & Ebsworth, M. (2018). The Use of Film Clips in a Viewing Time Task of Sexual Interests. *Archives of Sexual Behavior*, *47*(3), 627–635. https://doi.org/10.1007/s10508-017-1108-0

Lancry-Dayan, O. C., Nahari, T., Ben-Shakhar, G., & Pertzov, Y. (2018). Do You Know Him? Gaze Dynamics Toward Familiar Faces on a Concealed Information Test. *Journal of Applied Research in Memory and Cognition*, *7*(2), 291–302. https://doi.org/10.1016/j.jarmac.2018.01.011

Lang, P. J., Bradley, M. M., & Cuthbert, B. M. (1997). *International Affective Picture System (IAPS): Technical Manual and Affective Ratings*. *1*, 39–58.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390–423. https://doi.org/10.1037/0033-2909.126.3.390

Langlois, J. H., Ritter, J. M., Roggman, L. A., & Vaughn, L. S. (1991). Facial diversity and infant preferences for attractive faces. *Developmental Psychology*, *27*(1), 79–84. https://doi.org/10.1037/0012-1649.27.1.79

Lavy, E., & van den Hout, M. (1993). Selective attention evidenced by pictorial and linguistic stroop tasks. *Behavior Therapy*, *24*(4), 645–657. https://doi.org/10.1016/S0005-7894(05)80323-5

Leknes, S., Wessberg, J., Ellingsen, D.-M., Chelnokova, O., Olausson, H., & Laeng, B. (2013). Oxytocin enhances pupil dilation and sensitivity to 'hidden' emotional expressions. *Social Cognitive and Affective Neuroscience*, *8*(7), 741–749. https://doi.org/10.1093/scan/nss062

Levin, R. J., & van Berlo, W. (2004). Sexual arousal and orgasm in subjects who experience forced or non-consensual sexual stimulation – a review. *Journal of Clinical Forensic Medicine*, *11*(2), 82–88. https://doi.org/10.1016/j.jcfm.2003.10.008

Liao, H.-I., Yeh, S.-L., & Shimojo, S. (2011). Novelty vs. Familiarity Principles in Preference Decisions: Task-Context of Past Experience Matters. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00043

Lippa, R. A. (2012). Effects of Sex and Sexual Orientation on Self-Reported Attraction and Viewing Times to Images of Men and Women: Testing for Category Specificity. *Archives of Sexual Behavior*, *41*(1), 149–160. https://doi.org/10.1007/s10508-011-9898-y

Lippa, R. A., Patterson, T. M., & Marelich, W. D. (2010). Looking at and Longing for Male and Female "Swimsuit Models": Men Are Much More Category Specific Than Women. *Social Psychological and Personality Science*, *1*(3), 238–245. https://doi.org/10.1177/1948550609359814

Little, A. C., DeBruine, L. M., & Jones, B. C. (2014). Sex differences in attraction to familiar and unfamiliar opposite-sex faces: Men prefer novelty and women prefer familiarity. *Archives of Sexual Behavior*, *43*(5), 973–981. https://doi.org/10.1007/s10508-013-0120-2

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). *Facial attractiveness: Evolutionary based research | Philosophical Transactions of the Royal Society B: Biological Sciences*.

https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2010.0404?casa_token=GDTcGhO499IAAAAA%3A4yKMClJywmoMjQaIGBozsAAFkmuxtv6ZD1CIxcU9O-XnJoiF4tk21ET1Dmg_xrWqZveAuzQ91eYhlTo

Locher, P., Unger, R., Sociedade, P., & Wahl, J. (1993). At first glance: Accessibility of the physical attractiveness stereotype. *Sex Roles*, *28*(11), 729–743. https://doi.org/10.1007/BF00289990

Love, R. E., Sloan, L. R., & Schmidt, M. J. (1976). Viewing pornography and sex guilt: The priggish, the prudent, and the profligate. *Journal of Consulting and Clinical Psychology*, *44*(4), 624–629. https://doi.org/10.1037/0022-006X.44.4.624

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

Mahoney, J. M., & Strassberg, D. S. (1991). Voluntary control of male sexual arousal. *Archives of Sexual Behavior*, *20*(1), 1–16. https://doi.org/10.1007/BF01543003

Mannfolk, C., Liberg, B., Abé, C., & Rahm, C. (2023). Altered Neural and Behavioral Response to Sexually Implicit Stimuli During a Pictorial-Modified Stroop Task in Pedophilic Disorder. *Biological Psychiatry Global Open Science*, *3*(2), 292–300. https://doi.org/10.1016/j.bpsgos.2022.02.004

Marlowe, C. M., Schneider, S. L., & Nelson, C. E. (1996). Gender and attractiveness biases in hiring decisions: Are more experienced managers less biased? *Journal of Applied Psychology*, *81*(1), 11–21. https://doi.org/10.1037/0021-9010.81.1.11

Matthews, G., Middleton, W., Gilmartin, B., & Bullimore, M. A. (1991). Pupillary diameter and cognitive load. *Journal of Psychophysiology*, *5*(3), 265–271.

Matthews, O., Davies, A., Vigo, M., & Harper, S. (2020). Unobtrusive arousal detection on the web using pupillary response. *International Journal of Human-Computer Studies*, *136*, 102361. https://doi.org/10.1016/j.ijhcs.2019.09.003

McDougal, D. H., & Gamlin, P. D. (2015). *Autonomic control of the eye*. *1*(5), 439-473. http://doi.org/10.1002/cphy.c140014

McGarrigle, R., Dawes, P., Stewart, A. J., Kuchinsky, S. E., & Munro, K. J. (2017). Pupillometry reveals changes in physiological arousal during a sustained listening task. *Psychophysiology*, *54*(2), 193–203. https://doi.org/10.1111/psyp.12772

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the emotional Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(6), 1595–1607. https://doi.org/10.1037/0278-7393.21.6.1595

McPhail, I. V., Hermann, C. A., Fernane, S., Fernandez, Y. M., Nunes, K. L., & Cantor, J. M. (2019). *Validity in Phallometric Testing for Sexual Interests in Children: A Meta-Analytic Review*. *26*(3). https://doi.org/10.1177/1073191117706139

Meston, C. M., Heiman, J. R., Trapnell, P. D., & Paulhus, D. L. (1998). Socially desirable responding and sexuality self-reports. *The Journal of Sex Research*, *35*(2), 148–157. https://doi.org/10.1080/00224499809551928

Metzger, L. J., Orr, S. P., Lasko, N. B., McNally, R. J., & Pitman, R. K. (1997). Seeking the source of emotional Stroop interference effects in PTSD: A study of P3s to traumatic words. *Integrative Physiological and Behavioral Science*, *32*(1), 43–51. https://doi.org/10.1007/BF02688612

Mitterschiffthaler, M. T., Williams, S. C. R., Walsh, N. D., Cleare, A. J., Donaldson, C., Scott, J., & Fu, C. H. Y. (2008). Neural basis of the emotional Stroop interference effect in major depression. *Psychological Medicine*, *38*(2), 247–256. https://doi.org/10.1017/S0033291707001523

Mokros, A., Dombert, B., Osterheider, M., Zappalà, A., & Santtila, P. (2010). Assessment of Pedophilic Sexual Interest with an Attentional Choice Reaction Time Task. *Archives of Sexual Behavior*, *39*(5), 1081–1090. https://doi.org/10.1007/s10508-009-9530-6

Mokros, A., Gebhard, M., Heinz, V., Marschall, R. W., Nitschke, J., Glasgow, D. V., Gress, C. L. Z., & Laws, D. R. (2013). Computerized Assessment of Pedophilic Sexual Interest Through Self-Report and Viewing Time: Reliability, Validity, and Classification Accuracy of the Affinity Program. *Sexual Abuse*, *25*(3), 230–258. https://doi.org/10.1177/1079063212454550

Monin, B. (2003). The Warm Glow Heuristic: When Liking Leads to Familiarity. *Journal of Personality and Social Psychology*, *85*(6), 1035–1048. https://doi.org/10.1037/0022-3514.85.6.1035

Morandini, J. S., Veldre, A., Holcombe, A. O., Hsu, K., Lykins, A., Bailey, J. M., & Dar-Nimrod, I. (2019). Visual Attention to Sexual Stimuli in Mostly Heterosexuals. *Archives of Sexual Behavior*, *48*(5), 1371–1385. https://doi.org/10.1007/s10508-019-1419-4

Moreland, R. L., & Beach, S. R. (1992). Exposure effects in the classroom: The development of affinity among students. *Journal of Experimental Social Psychology*, *28*(3), 255–276. https://doi.org/10.1016/0022-1031(92)90055-O

Moreland, R. L., & Zajonc, R. B. (1982). Exposure effects in person perception: Familiarity, similarity, and attraction. *Journal of Experimental Social Psychology*, *18*(5), 395–415. https://doi.org/10.1016/0022-1031(82)90062-2

Müller, K., Curry, S., Ranger, R., Briken, P., Bradford, J., & Fedoroff, J. P. (2014). Changes in Sexual Arousal as Measured by Penile Plethysmography in Men with Pedophilic Sexual Interest. *The Journal of Sexual Medicine*, *11*(5), 1221–1229. https://doi.org/10.1111/jsm.12488

Müller J. R., Philiastides,. G., & Newsome, W. T. (2005). Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences, 102*(3), 524-529. https://doi.org/10.1073/pnas.0408311101

Murphy, J., Ipser, A., Gaigg, S. B., & Cook, R. (2015). Exemplar variance supports robust learning of facial identity. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 577–581. https://doi.org/10.1037/xhp0000049

Murphy, L., Ranger, R., Paul Fedoroff, J., Stewart, H., Gregg Dwyer, R., & Burke, W. (2015). Standardization of Penile Plethysmography Testing in Assessment of Problematic Sexual Interests. *The Journal of Sexual Medicine*, *12*(9), 1853–1861. https://doi.org/10.1111/jsm.12979

Naber, M., Frässle, S., Rutishauser, U., & Einhäuser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of Vision*, *13*(2), 11. https://doi.org/10.1167/13.2.11

Norton, M. I., Frost, J. H., & Ariely, D. (2007). Less is more: The lure of ambiguity, or why familiarity breeds contempt. *Journal of Personality and Social Psychology*, *92*(1), 97–105. https://doi.org/10.1037/0022-3514.92.1.97

Norton, M. I., Frost, J. H., & Ariely, D. (2013). Less is often more, but not always: Additional evidence that familiarity breeds contempt and a call for future research. *Journal of Personality and Social Psychology*, *105*(6), 921–923. https://doi.org/10.1037/a0034379

Ó Ciardha, C., Attard-Johnson, J., & Bindemann, M. (2018). Latency-Based and Psychophysiological Measures of Sexual Interest Show Convergent and Concurrent Validity. *Archives of Sexual Behavior*, *47*(3), 637–649. https://doi.org/10.1007/s10508-017-1133-z

Ó Ciardha, C., & Gormley, M. (2013). Measuring sexual interest using a pictorial modified Stroop task, a pictorial Implicit Association Test and a Choice Reaction Time task. *Journal of Sexual Aggression*, *19*(2), 158–170. https://doi.org/10.1080/13552600.2012.677486

Oberauer, K. (2019). Working Memory and Attention – A Conceptual Analysis and Review. *Journal of Cognition*, *2*(1), 36. https://doi.org/10.5334/joc.58

Oliva, M., & Anikin, A. (2018). Pupil dilation reflects the time course of emotion recognition in human vocalizations. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-23265-x

Olson, I. R., & Marshuetz, C. (2005). Facial Attractiveness Is Appraised in a Glance. *Emotion*, *5*(4), 498–502. https://doi.org/10.1037/1528-3542.5.4.498

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences*, *87*(1), 256–259. https://doi.org/10.1073/pnas.87.1.256

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, *59*(1), 185–198. https://doi.org/10.1016/S1071-5819(03)00017-X

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peskin, M., & Newell, F. N. (2004). Familiarity Breeds Attraction: Effects of Exposure on the Attractiveness of Typical and Distinctive Faces. *Perception*, *33*(2), 147–157. https://doi.org/10.1068/p5028

Peterson, Z. D., Janssen, E., & Laan, E. (2010). Women's Sexual Responses to Heterosexual and Lesbian Erotica: The Role of Stimulus Intensity, Affective Reaction, and Sexual History. *Archives of Sexual Behavior*, *39*(4), 880–897. https://doi.org/10.1007/s10508-009-9546-y

Petrak, J. (2002). Rape: History, myths, and reality. In *The trauma of sexual assault: Treatment, prevention and practice* (pp. 1–18).

Pezzoli, P., Babchishin, K., Pullman, L., & Seto, M. C. (2022). Viewing Time Measures of Sexual Interest and Sexual Offending Propensity: An Online Survey of Fathers. *Archives of Sexual Behavior*, *51*(8), 4097–4110. https://doi.org/10.1007/s10508-022-02324-5

Pratt, N., Willoughby, A., & Swick, D. (2011). Effects of Working Memory Load on Visual Selective Attention: Behavioral and Electrophysiological Evidence. *Frontiers in Human Neuroscience*, *5*, 57. https://www.frontiersin.org/articles/10.3389/fnhum.2011.00057

Prehn, K., Kazzer, P., Lischke, A., Heinrichs, M., Herpertz, S. C., & Domes, G. (2013). Effects of intranasal oxytocin on pupil dilation indicate increased salience of socioaffective stimuli. *Psychophysiology*, *50*(6), 528–537. https://doi.org/10.1111/psyp.12042

Price, S. A., Beech, A. R., Mitchell, I., & Humphreys, G. W. (2013). Measuring Deviant Sexual Interest Using the Emotional Stroop Task. *Criminal Justice and Behavior*, *40*(9), 970–987. https://doi.org/10.1177/0093854813476264

Price, S. A., Beech, A. R., Mitchell, I. J., & Humphreys, G. W. (2012). The promises and perils of the emotional Stroop task: A general review and considerations for use with forensic samples. *Journal of Sexual Aggression*, *18*(3), 253–268. https://doi.org/10.1080/13552600.2010.545149

Price, S. A., & Karl Hanson, R. (2007). A modified Stroop task with sexual offenders: Replication of a study. *Journal of Sexual Aggression*, *13*(3), 203–216. https://doi.org/10.1080/13552600701785505

Quinsey, V. L., Ketsetzis, M., Earls, C., & Karamanoukian, A. (1996). Viewing time as a measure of sexual interest. *Ethology and Sociobiology*, *17*(5), 341–354. https://doi.org/10.1016/S0162-3095(96)00060-X

Reimers S., & Stewart, N. (2007). Adobe Flash as a medium for online experimentation: A test of reaction time measurement capabilities. *Behaviour research methods, 39*(3), 365-370. https://doi.org/10.3758/BF03193004

Reis, H. T., Maniaci, M. R., Caprariello, P. A., Eastwick, P. W., & Finkel, E. J. (2011). Familiarity does indeed promote attraction in live interaction. *Journal of Personality and Social Psychology*, *101*(3), 557–570. https://doi.org/10.1037/a0022885

Rellini, A. H., McCall, K. M., Randall, P. K., & Meston, C. M. (2005). The relationship between women's subjective and physiological sexual arousal. *Psychophysiology*, *42*(1), 116–124. https://doi.org/10.1111/j.1469-8986.2005.00259.x

Rhodes, G., Halberstadt, J., Jeffery, L., & Palermo, R. (2005). The Attractiveness of Average Faces is Not a Generalized Mere Exposure Effect. *Social Cognition*, *23*(3), 205–217. https://doi.org/10.1521/soco.2005.23.3.205

Richards, A., French, C. C., Johnson, W., Naparstek, J., & Williams, J. (1992). Effects of mood manipulation and anxiety on performance of an emotional Stroop task. *British Journal of Psychology*, *83*(4), 479–491. https://doi.org/10.1111/j.2044-8295.1992.tb02454.x

Rieger, G., Cash, B. M., Merrill, S. M., Jones-Rounds, J., Dharmavaram, S. M., & Savin-Williams, R. C. (2015). Sexual arousal: The correspondence of eyes and genitals. *Biological Psychology*, *104*, 56–64. https://doi.org/10.1016/j.biopsycho.2014.11.009

Rieger, G., & Savin-Williams, R. C. (2012). The Eyes Have It: Sex and Sexual Orientation Differences in Pupil Dilation Patterns. *PLOS ONE*, *7*(8), e40256. https://doi.org/10.1371/journal.pone.0040256

Rieger, G., Savin-Williams, R. C., Chivers, M. L., & Bailey, J. M. (2016). Sexual arousal and masculinity-femininity of women. *Journal of Personality and Social Psychology*, *111*(2), 265–283. https://doi.org/10.1037/pspp0000077

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *Quarterly Journal of Experimental Psychology*, *70*(5), 897–905. https://doi.org/10.1080/17470218.2015.1136656

Ritchie, K. L., Palermo, R., & Rhodes, G. (2017). Forming impressions of facial attractiveness is mandatory. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-00526-9

Rönspies, J., Schmidt, A. F., Melnikova, A., Krumova, R., Zolfagari, A., & Banse, R. (2015). Indirect Measurement of Sexual Orientation: Comparison of the Implicit Relational Assessment Procedure, Viewing Time, and Choice Reaction Time Tasks. *Archives of Sexual Behavior*, *44*(5), 1483–1492. https://doi.org/10.1007/s10508-014-0473-1

Rosa, P. J. M., Oliveira, J. A. G., Alghazzawi, D., Fardoun, H., & Gamito, P. S. P. (2017). Affective and physiological correlates of the perception of unimodal and bimodal emotional stimuli. *Psicothema*. https://redined.educacion.gob.es/xmlui/handle/11162/149681

Rosenzweig, S. (1942). THE PHOTOSCOPE AS AN OBJECTIVE DEVICE FOR EVALUATING SEXUAL INTEREST*. *Psychosomatic Medicine*, *4*(2), 150.

Rullo, J. E., Strassberg, D. S., & Miner, M. H. (2015). Gender-Specificity in Sexual Interest in Bisexual Men and Women. *Archives of Sexual Behavior*, *44*(5), 1449–1457. https://doi.org/10.1007/s10508-014-0415-y

Rupp, H. A., & Wallen, K. (2007). Sex differences in viewing sexual stimuli: An eye-tracking study in men and women. *Hormones and Behavior*, *51*(4), 524–533. https://doi.org/10.1016/j.yhbeh.2007.01.008

Saegert, S., Swap, W., & Zajonc, R. B. (1973). Exposure, context, and interpersonal attraction. *Journal of Personality and Social Psychology*, *25*(2), 234–242. https://doi.org/10.1037/h0033965

Sanbonmatsu, D. M., Mazur, D., Pfeiffer, B. E., Kardes, F. R., & Posavac, S. S. (2012). The Less the Public Knows the Better? The Effects of Increased

Knowledge on Celebrity Evaluations. *Basic and Applied Social Psychology*, *34*(6), 499–507. https://doi.org/10.1080/01973533.2012.728408

Santtila, P., Mokros, A., Viljanen, K., Koivisto, M., Sandnabba, N. K., Zappalà, A., & Osterheider, M. (2009). Assessment of sexual interest using a choice reaction time task and priming: A feasibility study. *Legal and Criminological Psychology*, *14*(1), 65–82. https://doi.org/10.1348/135532507X267040

Schmidt, A. F., Babchishin, K. M., & Lehmann, R. J. B. (2017). A Meta-Analysis of Viewing Time Measures of Sexual Interest in Children. *Archives of Sexual Behavior*, *46*(1), 287–300. https://doi.org/10.1007/s10508-016-0806-3

Schneider,. A., & Kastner, S. (2009). Effects of sustained spatial attention in the human lateral geniculate nucleus and superior colliculus. *Journal of Neuroscience, 29*(6), 1784-1795. https://doi.org/10.1523/JNEUROSCI.4452-08.2009

Schubert, T. W., Murteira, C., Collins, E. C., & Lopes, D. (2013). ScriptingRT: A software library for collecting response latencies in online studies of cognition. *PloS One, 8*(6), e67769. https://doi.org/10.1371/journal.pone.0067769

Schwartz, R., Rothermich, K., Kotz, S. A., & Pell, M. D. (2018). Unaltered emotional experience in Parkinson's disease: Pupillometry and behavioral evidence. *Journal of Clinical and Experimental Neuropsychology*, *40*(3), 303–316. https://doi.org/10.1080/13803395.2017.1343802

Seto, M. C., Harris, G. T., Rice, M. E., & Barbaree, H. E. (2004). The Screening Scale for Pedophilic Interests Predicts Recidivism Among Adult Sex Offenders with Child Victims. *Archives of Sexual Behavior*, *33*(5), 455–466. https://doi.org/10.1023/B:ASEB.0000037426.55935.9c

Siegle, G. J., Steinhauer, S. R., Carter, C. S., Ramel, W., & Thase, M. E. (2003). Do the Seconds Turn Into Hours? Relationships between Sustained Pupil Dilation in Response to Emotional Information and Self-Reported Rumination. *Cognitive Therapy and Research*, *27*(3), 365–382. https://doi.org/10.1023/A:1023974602357

Slater, A., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., & Samuels, C. (1998). Newborn infants prefer attractive faces. *Infant Behavior and Development*, *21*(2), 345–354. https://doi.org/10.1016/S0163-6383(98)90011-X

Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., Carlson, J. M., & Schooler, J. W. (2011). Pupillometric Evidence for the Decoupling of Attention from Perceptual Input during Offline Thought. *PLOS ONE*, *6*(3), e18298. https://doi.org/10.1371/journal.pone.0018298

Smith, P., & Waterman, M. (2004). Processing Bias for Sexual Material: The Emotional Stroop and Sexual Offenders. *Sexual Abuse: A Journal of Research and Treatment*, *16*(2), 163–171. https://doi.org/10.1023/B:SEBU.0000023064.54397.d6

Snowden, R. J., McKinnon, A., Fitoussi, J., & Gray, N. S. (2019). Pupillary Responses to Static Images of Men and Women: A Possible Measure of Sexual Interest? *The Journal of Sex Research*, *56*(1), 74–84. https://doi.org/10.1080/00224499.2017.1394959

Snowden, R. J., O'Farrell, K. R., Burley, D., Erichsen, J. T., Newton, N. V., & Gray, N. S. (2016). The pupil's response to affective pictures: Role of image duration, habituation, and viewing mode. *Psychophysiology*, *53*(8), 1217–1223. https://doi.org/10.1111/psyp.12668

Song, F., Hooper, L., & Loke, Y. K. (2013). Publication bias: What is it? How do we measure it? How do we avoid it? *Open Access Journal of Clinical Trials*, *5*, 71–81. https://doi.org/10.2147/OAJCT.S34419

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, Involuntary Top-Down Guidance of Attention From Working Memory. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(2), 248–261. https://doi.org/10.1037/0096-1523.31.2.248

South Palomares, J. K., & Young, A. W. (2018). Facial First Impressions of Partner Preference Traits: Trustworthiness, Status, and Attractiveness. *Social Psychological and Personality Science*, *9*(8), 990–1000. https://doi.org/10.1177/1948550617732388

Spada, A. H., & Jeglic, E. L. (2016). A cognitive-based indicator of deviant sexual interest: Concurrent validation of the Stroop task. *Journal of Sexual Aggression*, *22*(2), 246–262. https://doi.org/10.1080/13552600.2015.1078000

Sprecher, S., Treger, S., & Wondra, J. D. (2013). Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions. *Journal of Social and Personal Relationships*, *30*(4), 497–514. https://doi.org/10.1177/0265407512459033

Steidtmann, D., Ingram, R. E., & Siegle, G. J. (2010). Pupil response to negative emotional information in individuals at risk for depression. *Cognition and Emotion*, *24*(3), 480–496. https://doi.org/10.1080/02699930902738897

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Suschinsky, K. D., & Lalumière, M. L. (2011). *Prepared for Anything?: An Investigation of Female Genital Arousal in Response to Rape Cues*. *22*(2), 159–165. https://doi.org/10.1177/0956797610394660

Suschinsky, K. D., Lalumière, M. L., & Chivers, M. L. (2009). Sex Differences in Patterns of Genital Sexual Arousal: Measurement Artifacts or True Phenomena? *Archives of Sexual Behavior*, *38*(4), 559–573. https://doi.org/10.1007/s10508-008-9339-8

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*, *108*(2), 397–415. https://doi.org/10.1111/bjop.12206

Tanner, J. M. (1973). Growing Up. *Scientific American*, *229*(3), 34–43.

Taylor, S. F., Kornblum, S., Minoshima, S., Oliver, L. M., & Koeppe, R. A. (1994). Changes in medial cortical blood flow with a stimulus-response compatibility task. *Neuropsychologia*, *32*(2), 249–255. https://doi.org/10.1016/0028-3932(94)90010-8

Thiruchselvam, R., Harper, J., & Homer, A. L. (2016). Beauty is in the belief of the beholder: Cognitive influences on the neural response to facial attractiveness. *Social Cognitive and Affective Neuroscience*, *11*(12), 1999–2008. https://doi.org/10.1093/scan/nsw115

Thornhill, R., & Palmer, C. T. (2001). *A Natural History of Rape: Biological Bases of Sexual Coercion*. MIT Press.

Thornton, D., & Laws, D. R. (2009). *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders*. https://doi.org/10.1002/9780470747551

Trottier, D., Rouleau, J.-L., Renaud, P., & Goyette, M. (2014). Using Eye Tracking to Identify Faking Attempts During Penile Plethysmography Assessment. *The Journal of Sex Research*, *51*(8), 946–955. https://doi.org/10.1080/00224499.2013.832133

Ullrich, J., Krueger, J. I., Brod, A., & Groschupf, F. (2013). More is not less: Greater information quantity does not diminish liking. *Journal of Personality and Social Psychology*, *105*(6), 909–920. https://doi.org/10.1037/a0033183

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior Cingulate Cortex, Conflict Monitoring, and Levels of Processing. *NeuroImage*, *14*(6), 1302–1308. https://doi.org/10.1006/nimg.2001.0923

Võ, M. L.-H., Jacobs, A. M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology*, *45*(1), 130–140. https://doi.org/10.1111/j.1469-8986.2007.00606.x

Vogels, J., Demberg, V., & Kray, J. (2018). The Index of Cognitive Activity as a Measure of Cognitive Processing Load in Dual Task Settings. *Frontiers in Psychology*, *9*. https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02276

Wang, C.-A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., & Munoz, D. P. (2018). Arousal Effects on Pupil Size, Heart Rate, and Skin Conductance in an Emotional Face Task. *Frontiers in Neurology*, *9*. https://www.frontiersin.org/articles/10.3389/fneur.2018.01029

Watson, A. B., & Yellott, J. I. (2012). A unified formula for light-adapted pupil size. *Journal of Vision*, *12*(10), 12. https://doi.org/10.1167/12.10.12

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology*, *77*(1), 97–108. https://doi.org/10.1111/j.2044-8295.1986.tb01985.x

Watts, T. M., Holmes, L., Savin-Williams, R. C., & Rieger, G. (2017). Pupil Dilation to Explicit and Non-Explicit Sexual Stimuli. *Archives of Sexual Behavior*, *46*(1), 155–165. https://doi.org/10.1007/s10508-016-0801-8

Weiss, M. W., Trehub, S. E., Schellenberg, E. G., & Habashi, P. (2016). Pupils dilate for vocal or familiar music. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1061–1065. https://doi.org/10.1037/xhp0000226

Wells, B. M., Skowronski, J. J., Crawford, M. T., Scherer, C. R., & Carlston, D. E. (2011). Inference making and linking both require thinking: Spontaneous trait inference and spontaneous trait transference both rely on working memory capacity. *Journal of Experimental Social Psychology*, *47*(6), 1116–1126. https://doi.org/10.1016/j.jesp.2011.05.013

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*(1), 3–24. https://doi.org/10.1037/0033-2909.120.1.3

Williams, J. M. G., & Nulty, D. D. (1986). Construct accessibility, depression and the emotional stroop task: Transient mood or stable structure? *Personality and Individual Differences*, *7*(4), 485–491. https://doi.org/10.1016/0191-8869(86)90127-3

Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Wingenfeld, K., Mensebach, C., Driessen, M., Bullig, R., Hartje, W., & Beblo, T. (2006). Attention Bias towards Personally Relevant Stimuli: The Individual Emotional Stroop Task. *Psychological Reports*, *99*(3), 781–793. https://doi.org/10.2466/PR0.99.3.781-793

Wolchik, S. A., Braver, S. L., & Jensen, K. (1985). Volunteer bias in erotica research: Effects of intrusiveness of measure and sexual background. *Archives of Sexual Behavior*, *14*(2), 93–107. https://doi.org/10.1007/BF01541656

Worling, J. R. (2006). Assessing Sexual Arousal with Adolescent Males Who Have Offended Sexually: Self-Report and Unobtrusively Measured Viewing Time. *Sexual Abuse*, *18*(4), 383–400. https://doi.org/10.1177/107906320601800406

Wright, L. W., & Adams, H. E. (1999). The effects of stimuli that vary in erotic content on cognitive processes. *The Journal of Sex Research*, *36*(2), 145–151. https://doi.org/10.1080/00224499909551979

Xu, Y., Rahman, Q., & Zheng, Y. (2017). Gender-Specificity in Viewing Time Among Heterosexual Women. *Archives of Sexual Behavior*, *46*(5), 1361–1374. https://doi.org/10.1007/s10508-016-0795-2

Xue, S.-W., Wu, H.-B., Zhang, L., & Zhang, D.-X. (2020). Intranasal Oxytocin Increases Perceptual Salience of Faces in the Absence of Awareness. *Psychiatry Investigation*, *17*(4), 292–298. https://doi.org/10.30773/pi.2019.0130

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, *51*(3), 277–284. https://doi.org/10.1111/psyp.12151

Zele, A. J., & Gamlin, P. D. (2020). Editorial: The Pupil: Behavior, Anatomy, Physiology and Clinical Biomarkers. *Frontiers in Neurology*, *11*. https://www.frontiersin.org/articles/10.3389/fneur.2020.00211

Zhou, J., Xu, T., Zhao, X., Chen, J., Zhao, Y., Dong, H., & Zhang, M. (2021). The automatic processing of social information in working memory. *Journal of Experimental Psychology: General*, *150*(11), 2321–2345. https://doi.org/10.1037/xge0001046