Lighting for Pedestrians’ Interpersonal Evaluations: Is the Face the Critical Cue?

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

By:

Khalid HAMOODH

BArch, MID

The University of Sheffield

School of Architecture

Submission Date

August 2022

Abstract

Road lighting in minor roads is designed primarily for the needs of pedestrians and cyclists. Pedestrians' considerations include their safety and their feeling of safety after-dark. One factor that contributes to a pedestrian's feeling of safety is the ability to evaluate other people (known as interpersonal evaluations); road lighting should enhance the ability to make the visual component of this evaluation. While past lighting studies have assumed the face is the critical target, and hence investigated the effect of lighting on facial recognition, this assumption has yet to be verified. Different personal features subtend targets of different sizes, colours, and contrasts, so a better understanding of the important visual cues is essential for appropriate research.

Three experiments were conducted to investigate the visual cues used by pedestrians in interpersonal evaluations. The first two used a subjective evaluation procedure, in which images portraying pedestrians in different situations were evaluated using either category rating or both category rating and paired comparisons. The assumption was that those situations rated (or chosen) as less safe would indicate critical visual cues. Experiment 1 compared the relative importance of gender, number of people, walking direction, direction of the fall of light, and the exposure of the face and hands: the results suggested that the exposure of face and hands had a greater effect than other cues, but did not distinguish between the face and hands. Experiment 2 was therefore conducted to further investigate exposure to view of the face and hands: the results suggested that the face is a more important cue than the hands. To test the robustness of these findings, Experiment 3 investigated the same question as Experiments 1 and 2 but used an objective measure, eye-tracking. The core assumption was that the most important visual cue in a scene is that which receives the first visual fixation and the longest duration of visual fixations: the results again suggested that the face is the most important visual cue in interpersonal evaluations.

It is therefore concluded that the face is the key visual cue used in pedestrians' interpersonal evaluations. This supports the assumption of previous research investigating the effect of changes in lighting for pedestrians.

Acknowledgements

First and foremost, my thanks go to my god (Allah) for giving me blessings, strength, and forming the circumstances that allowed me to complete my entire PhD journey.

To my country, the Kingdom of Saudi Arabia, for funding my scholarship, and for helping me become all that I am today.

To my supervisor, Prof. Steve Fotios; I am grateful to him for giving me the opportunity to pursue my education. I could not have asked for a better supervisor for my PhD.

To my parents and parents-in-law who believed in me, and were always there during the hard times. Thanks for providing me with unconditional care and love.

To my dearest wife and my best friend Raghda Zabeedi, I am so blessed for her support during all my studies. I am grateful for her patience and the sacrifices she made to support me with continuous encouragement. She was, and will always be, the irreplaceable person.

To my lovely sons Majdi and Yousef for bringing joy into my world and making me smile no matter what. They have witnessed this entire journey and shared life's difficulties since they were two and four years old. I would like to say SORRY for being grumpy whilst I wrote this thesis, I LOVE you both very much,

To my siblings for their unconditional support, I am so grateful that we are so close.

To my colleagues in the lighting research group, including Dr Chris Cheal for building the experimental apparatus and software programs; Dr James Uttley for his kindness, feedback, and help with data/statistical analysis; Dr Yichong Mao, Dr Aleksandra Monteiro, Intisar Husain, and Scott Fox; and special thanks to Dr Hussain Qasem who taught me how to start using the eye-tracking device and provided me with invaluable advice and psychological support, enabling me to pursue my research with no worries.

To my colleagues at the University of Sheffield who agreed to be photographed and be part of this study, especially Tha'er Abdalla who provided me with psychological and physical support to overcome my academic anxiety; I wish him a successful thesis defence soon.

Finally, to all individuals who helped me directly or indirectly during this unforgettable journey.

Papers Related to this Thesis

Journal paper

Hamoodh, K., Fotios, S., & Cheal, C., (2022). Visual cues to interpersonal evaluations for pedestrians. *Lighting Research & Technology*. https://doi.org/10.1177/14771535221093470

Fotios, S., Mao, Y., **Hamoodh**, K., & Cheal, C. (2022). Using relative visual performance to predict performance of an interpersonal evaluation task with variation in adaptation luminance, observer age, skin tone, pavement reflection and interpersonal distance. *Lighting Research & Technology*. https://doi.org/10.1177/14771535211069027

Conference proceedings papers

Hamoodh, K., & Fotios, S. (2021). The visual cues used to evaluate other pedestrians; face, body, or something else? *In Proceedings of the Conference CIE 2021: Light for Life - Living with Light, hosted by NC Malaysia online,* Sep 27 – 29, 2021, pp.519-524. DOI 10.25039/x48.2021.OP64

Fotios, S., **Hamoodh**, K., & Clanton, N. (2019). What are you looking at? Testing Nancy's rules for pedestrian interactions. *In Proceedings of the 29th Quadrennial Session of the CIE in Washington D.C., USA,* June 14 – 22, 2019, pp.1669-1674. DOI 10.25039/x46.2019.PO169

Contents

List of Figures

List of Tables

Chapter 1. Road Lighting for Pedestrians

1.1. Introduction

Imagine that you are walking alone, after-dark, and notice another person ahead: do you choose to keep walking or instead take action to avoid them? One aim of road lighting is to support such an evaluation (BS 5489-1:2020) by revealing important visual cue(s) about other pedestrians. Lighting research should, therefore, investigate how the visibilities of such cues are influenced by changes in lighting characteristics. Thus, the first question for such lighting research is, what is/are the important visual cue(s) in evaluating other pedestrians?

This work concerns road lighting and walking after-dark. To begin with, walking has benefits for personal health and fitness: it promotes well-being, mental alertness, energy, and a positive mood (Ekkekakis et al., 2000), as well as physical effects such as losing weight, improving cholesterol, and controlling hypertension (Rippe et al., 1988). Walking is a simple, easy, free exercise, and mode of transport. People can walk almost anywhere without the need for equipment or expenses such as a gym membership. However, after-dark, people feel a lower degree of safety (Fisher & Nasar, 1992; Gover et al., 2011), especially in dark places and poorly lit areas which are often seen as dangerous (Nasar & Fisher, 1992; Nasar & Jones, 1997). Previous studies have highlighted the effect of lighting as the most important aspect of the physical environment in perceived personal safety (Blöbaum & Hunecke, 2005; Loewen et al., 1993; Nasar et al., 1993; Nasar & Jones, 1997).

In qualitative interviews, people stated that they hesitate to leave their homes because they are afraid of darkness, and they feel safer when there is light (Bensch et al., 2012). Road lighting helps illuminate outdoor places and offsets the absence of daylight, and numerous studies have found that a higher level of light improves the sense of being safe (Blöbaum & Hunecke, 2005; Boomsma & Steg, 2014; Boyce et al., 2000; Peña-García et al., 2015; Vrij & Winkel, 1991). Considering this, road lighting encourages people to go after-dark.

The International Commission on Illumination (CIE) is a worldwide cooperation between lighting researchers to discuss and exchange all lighting related matters. The CIE 115:2010 listed the main purposes of installing road lighting to be the following: (1) to allow road users to proceed safely; (2) to allow pedestrians to see hazards, orientate themselves, recognise other pedestrians, and give them a sense of security; (3) to improve the appearance of the environment, at both day and night.

In the British Standard BS 5489-1:2020, the main purposes of installing road lighting in subsidiary roads are stated to:

1. Assist pedestrians and cyclists to orientate themselves and detect other vehicle and hazards.
2. Allow pedestrians to recognise other pedestrians and feel more secure; where recognition and personal safety were described as visual tasks undertaken by pedestrians to be safe and feel safe with respect to two sets of factors:
    A. Environmental factors in an appropriately lit street: visual comfort, and perceived ability to judge the intent and/or identity of other road users.
    B. Movement factors: trip hazard, and to judge the intent and/or identity of other people.
3. Reduce crime and fear of crime against people and property.
4. Improve the appearance of the environment, at both day and night.
5. Provide some guidance for motorists.

The current research focuses on purpose number 2 from each of the CIE 115:2010 and BS 5489-1:2020 which is the task of evaluating other pedestrians (or judging the intent and/or identity of other people) as this affects the sense of safety felt by pedestrians. This chapter introduces and defines road lighting for pedestrians on subsidiary roads. It highlights their needs when walking after-dark, focusing on their visual task of conducting interpersonal evaluations (also known as interpersonal/social judgements). It also reviews the standard of road lighting installed in the UK to see if it meets those needs.

1.2. Road lighting in subsidiary roads

Road lighting is the system of components installed to illuminate public spaces such as roads, footpaths, cycle paths, parks, and pedestrian areas. Typically, road lighting in the UK is provided by luminaires mounted on top of 6 m high lamp-posts, which are spaced at intervals of 20 – 30 m. This thesis focuses on subsidiary roads (also known as minor or residential roads), and this also applies to pedestrian areas, footpaths, and cycle tracks. Figure 1.1 depicts an example of this road lighting after-dark.

Figure 1.1. Example of road lighting after-dark in Sheffield, UK (Photograph by author).

Light in general – and road lighting in particular – has enormous benefits for people's lives such as enhancing security, providing safety, and facilitating safe movement by increasing the visibility of the road ahead. People appreciate these benefits and they will not give it up easily (Boyce, 2019).

Human vision needs light and fails in complete darkness. Surfaces are seen because they reflect light towards an observer's eye. In daylight, we might experience luminance of around 500 cd/m^2 (looking towards grass), whereas after-dark this reduces to around 1.0 cd/m^2 (looking towards a concrete road surface) (Boyce, 2014, p. 9, Table 1.3). For an eye to detect an object, a visual contrast threshold is needed which is the minimum contrast generated by a given object that makes the object visible against a given background (CIE, 2016, p. 155). For a target subtending 1 min arc at the eye (considered the visual acuity corresponding to normal vision), a contrast threshold of 2 requires an adaptation luminance of 1 cd/m^2 (Boyce, 2014, p. 65, 68, Figure 2.15) whereas for a larger object (e.g. subtending 1 degree of arc) a the threshold contrast reduces to about 0.05 at 1 cd/m^2 – it is easier to see larger objects. This means that we need a higher target-to-background contrast to see something after-dark than we do in daylight.

Similarly, acuity (the ability to see small details, commonly characterised by the gap size of the smallest visible Landolt ring) reduces from 2 min arc^{-1} at 500 cd/m^2 to 0.4 at 1 cd/m^2 (Boyce, 2014, p. 70, Figure 2.17). This means that in daylight we can see smaller details than we can after-dark. Thus, after-dark, low light levels mean visual performance is impaired.

After-dark, when the sun has set, the light condition is dim and the natural source of light is sunlight reflected from the moon. In this level of darkness, people might have a greater fear of threats, probably because in darkness it is harder to detect potential dangers (Schaller, Park, & Mueller, 2003; Veenstra & Koole, 2018). Thus, the dark areas give pedestrians a feeling of insecurity (Simons et al., 1987), or feel less safe (Fotios, Unwin, et al., 2015), Therefore, people who do not anticipate feeling safe are less likely to walk after-dark (Foster et al., 2004; Roman & Chalfin, 2008).

In addition, after-dark, pedestrians want to be able to walk safely, to see where they are, and to appreciate their surroundings (Boyce, 2014, p. 427). To walk safely includes physical safety of being safe and mental safety of feeling safe. To be safe means to be able to detect potential hazards (BS 5489-1:2013), and to feel safe means to see the way ahead and if there is any potential escape route (Fisher and Nasar, 1992). After-dark, road lighting aims to achieve these kinds of safety by helping pedestrians to see the environment and other people.

Hence, the UK government invests in modernising street lighting, and this expenditure has increased annually from £20 million in 2016 to £28 million in 2018 and £54 million in 2021 (HM Treasury, 2021). Thus, road lighting can provide (or at least is expected to provide) a safeguard for pedestrians (Rea, 2000) and increases their feeling of safety (Van Cauwenberg et al., 2012). Therefore, road lighting designers have a duty to provide safe walking environments to decrease pedestrians' fear of crimes and encounters with other people.

The design of road lighting in England and Wales tends to follow the recommendations of the British Standard BS 5489-1:2020. This classifies road lighting according to four different areas: traffic routes, conflict areas, subsidiary roads, and city and town centres. There are separate considerations and recommendations for lighting in each type of area, but in general, they focus on providing sufficient lighting for a specific area, these depending on the visual tasks required. For subsidiary roads, there are four classes of lighting, P3 to P6. Table 1.1 shows how these classes are chosen according to traffic flow and ambient luminance. It is important to note that while BS 5489-1:2020 indicates what lighting class to use in a given situation (Table 1.1), the lighting conditions for that class are given in BS EN 13201-2:2015 (Table 1.2).

Table 1.1. Lighting classes for subsidiary roads with slow-moving vehicles, cyclists, and pedestrians. This is Table A.5 of BS 5489-1:2020.

Traffic flow	**Lighting class**		
	E1 to E4	**E1 to E2**	**E3 to E4**
	Pedestrian and cyclist only	**Speed limit v ≤ 30 mph**	**Speed limit v ≤ 30 mph**
Busy	P5	P4	P3
Normal	P5	P5	P4
Quiet	P6	P5	P4

Table 1.2. P lighting classes. This is Table 3 of BS EN 13201-2:2015.

Class	**Horizontal illuminance**		**Additional requirement if facial recognition is necessary**	
	$\bar{E}$ [a] **[minimum maintained] lx**	E_{min} **[maintained] lx**	$E_{V,min}$ **[maintained] lx**	$E_{SC,min}$ **[maintained] lx**
P3	7.50	1.50	2.5	1.5
P4	5.00	1.00	1.5	1.0
P5	3.00	0.60	1.0	0.6
P6	2.00	0.40	0.6	0.2

[a] To provide for uniformity, the actual value of the maintained average illuminance shall not exceed 1.5 times the minimum $\bar{E}$ value indicated for the class.

For subsidiary roads, the primary quantity is defined as the minimum average illuminance on the pavement surface. However, these illuminances were proposed by a committee based on their experience; the basis of their decisions was not reported (Fotios & Gibbons, 2018). The aim of the current work is to contribute towards establishing an empirical basis for road lighting, one expected benefit for pedestrians is to aid the interpersonal evaluation of other people after-dark. However, before the effect of changes in lighting can be investigated, there is a need to establish the important visual cues.

In summary, road lighting on subsidiary roads is designed primarily for the safety needs of pedestrians and cyclists, enabling detecting hazards and to discourage crime against them (BS 5489-1:2020). Vision plays a role in all of these needs, especially after-dark when vision deteriorates (Plainis et al., 2005), thus, road lighting is installed to offset this and help maintain the flow of visual information.

1.3. Evaluating other pedestrians

Regarding the needs of pedestrians, Caminada and Van Bommel (1984) suggested three requirements: first, to detect details of their surroundings, for instance potential trip hazards; second, to examine other users closely, namely "identification of persons or of intentions"; and third, to generally feel safe. This thesis focuses only on the second requirement: evaluating the intentions of other people, also known as interpersonal evaluation (or interpersonal/social judgment). This evaluation is essential to avoid potential threat but what visual cues are used when evaluating other people? Knowing this will help to define what should be lit and hence establish design criteria for road lighting. To achieve this, it is important to understand what people tend to look at and where.

When pedestrians are walking along a road after-dark on their own and see another pedestrian ahead, what visual cues about the approaching person tell them whether it is safe to continue or whether an avoidance action should be taken? Road lighting plays an important role in this interpersonal evaluation because the effect of perceived lighting quality determines people's decision either to walk or to avoid a certain path (Johansson & Rahm, 2015).

Evaluations of other people may be informed by cues. There are many general types of cues that a pedestrian could use for evaluating their personal safety. These may be categorised as visual and non-visual (or, verbal and non-verbal). Examples include the cues available from the presence of other people, animals and the appearance of the immediate environment. Among these cues, road lighting is not expected to affect the non-visual cues such as those associated with sound.

The visual cues may be informed by body posture, eye contact, and facial recognition and expression, as well as the gender, number, and direction of travel of other people, direction of light on those approaching, and the visibility of their faces and their hands

(Fotios & Johansson, 2019; Willis et al., 2011b, 2011a). We do not know how important each of these cues is, if at all, nor the hierarchy of these potential visual cues in relation to intent.

Road lighting standards tend to refer to facial recognition as an identification measure in interpersonal evaluations; BS 5489-1:2013 states *“Risk of crime or need for facial recognition: where there is a need for better facial recognition”;* BS EN 13201-2:2015 refers to facial recognition (e.g., Table 3 in that document specifies “*Additional requirement if facial recognition is necessary*”); BS 5489-1:2020 states *“Vertical illuminance is important in car parks, for facial recognition, personal security and CCTV”*; and CIE 136:2000 states *“It is important that such intentions be determined, usually from facial expressions, not less than 4 m away to enable the viewer to take any necessary avoiding action”.* However, this was suggested to be an incorrect definition of the evaluation made when encountering others because (1) recognition of identity does not generally say anything about intention, and (2) if the person is unknown, no amount of light will reveal their identity (Fotios & Johansson, 2019).

If the approaching person is a known/familiar, the face could be an incorrect target for evaluating the effect of changes in lighting. This is because (1) familiar faces are easy to identify because they are memorised better than unfamiliar faces (Chapman et al., 2018), (2) familiar faces can be processed and located faster than unfamiliar faces (Persike et al., 2013), and (3) familiar faces can be easily identified, even from very low quality images (Burton et al., 1999), or if the face is inverted, negative, or has been bottom-lit only (Johnston et al., 1992), or if it is severely distorted (Hole et al., 2002).

In lighting research, the interpersonal evaluation was investigated based on two basis of facial recognition: Facial Identity Recognition (FIR) and Facial Emotion Recognition (FER). To clarify, FIR is for a known person only (by giving their name) using facial properties such as eyes, mouth, and nose to identify the person. This is achieved by face neurons in the amygdala (Rutishauser et al., 2011). Conversely, FER, also known as the identification of emotion as portrayed by facial expression, or facial expressions of emotion; FER defines the ability to discriminate between facial expressions using codes from changeable face parts such as eye gaze and lip movement. This is conducted by a dedicated system which analyses the shape of the face in making an expression such as happy, sad, disgust…etc (Bruce & Young, 1986). Thus, FER gives clues as to an individual's feelings, and is used for both familiar and unfamiliar persons.

Until recently, the majority of past lighting studies investigated interpersonal evaluation on FIR (Alferdinck et al., 2010; Boyce & Rea, 1990; Caminada & Van Bommel, 1984; Dong et al., 2015; Iwata et al., 2015; Johansson et al., 2001; Knight, 2010; Knight et al., 2007; Lin & Fotios, 2015; Okud & Satoh, 2000; Raynham & Saksvikronning, 2003; Rea et al., 2009; Rombauts et al., 1989; Romnée & Bodart, 2014; Yao et al., 2009) and several investigations have been conducted around FER (Fotios et al., 2015, 2017, 2018; Johansson & Rahm, 2015; Li & Yang, 2018; Rahm & Johansson, 2018; Yang & Fotios, 2015). However, FIR does not say anything about intentions; in terms of feeling safe, it is likely the intentions of other people that matter, not who they are. On the other hands, we can estimate peoples' intentions from the FER (described as a potential inference of people's internal state) (Knutson, 1996).

Fotios & Johansson (2019) suggested FER is a better way to evaluate the intentions of other people; the expressions of other people were suggested to be the correct focal point in determining whether we should feel safe. Therefore, FIR may not be the correct, or only, visual cue, and FER suggested to be the correct. Recent research has focused on FER instead of FIR. One reason for this change is because FER is associated with approach-avoid decisions relating to unfamiliar people (Willis et al., 2011a). However, FER still uses the face as the target. Whether it is FIR or FER that matters, further research is needed to confirm that the face is the correct (and most important) visual cue for interpersonal evaluation, where much previous work assumes the face is the critical target, but this assumption is yet to be tested.

1.4. Research aims

As previously stated, road lighting plays a critical role in pedestrians' safety and feelings of safety after-dark, and is one of the main goals of installing road lighting. As such, the aim of this thesis is to investigate the visual cues used in interpersonal evaluations of other pedestrians when walking after-dark, and, if possible, to identify which of these cues are the most important. To feel safe after-dark, pedestrians need to see information about the approaching person. Thus, a critical visual task to help them feel safe is that interpersonal evaluation.

Interpersonal evaluation uses visual cues about the approaching person such as gender, face, clothes, hands, and relative direction of travel. It is not yet known which is/are the most important visual cue(s). While it is widely assumed that the face is the most important cue, this has not yet been verified through empirical evidence. Establishing the critical visual cue(s) will support ongoing research on optimal lighting for interpersonal evaluations.

This thesis discusses the potentially pertinent variables for establishing a hierarchy of factors that make people feel safe and be safe when they walk after-dark in the street. To this purpose, three experiments were conducted. The first explored potential visual cues; the second focused on the most influential cue found in the first experiment. The third experiment addressed the same question but employing a different method in order to generate robust results. The results of this research will contribute to ongoing research on current UK lighting standards for residential roads in order to make recommendations for development for the British Standards Institution (BS EN).

1.5. Structure of this thesis

Chapter 2 presents a literature review of interpersonal evaluations and formulates the research questions. The first two experiments were conducted to compare evaluations of safety when encountering different pedestrians; Chapter 3 describes the method, and the results of which are presented in Chapters 4 and 5. These both experiments used subjective evaluations, hence, a third experiment was conducted to validate their findings using an alternative experimental design with an objective measure – eye-tracking. The method and results for this experiment are presented in Chapters 6 and 7, respectively. Chapter 8 discusses the thesis as a whole and considers the limitations of the experiments. Finally, the conclusions of this study and recommendations for further work are presented in Chapter 9. This structure is summarised in Figure 1.2.

This thesis used empirical experimentation as its primary methodology. The data were collected from human participants and analysed using quantitative techniques. The methodology investigated potential visual cues used for interpersonal evaluation after-dark. The implications of the results are applicable to all people, as they share a similar physiological mechanism regardless of cultural differences or an individual's history of personal experiences.

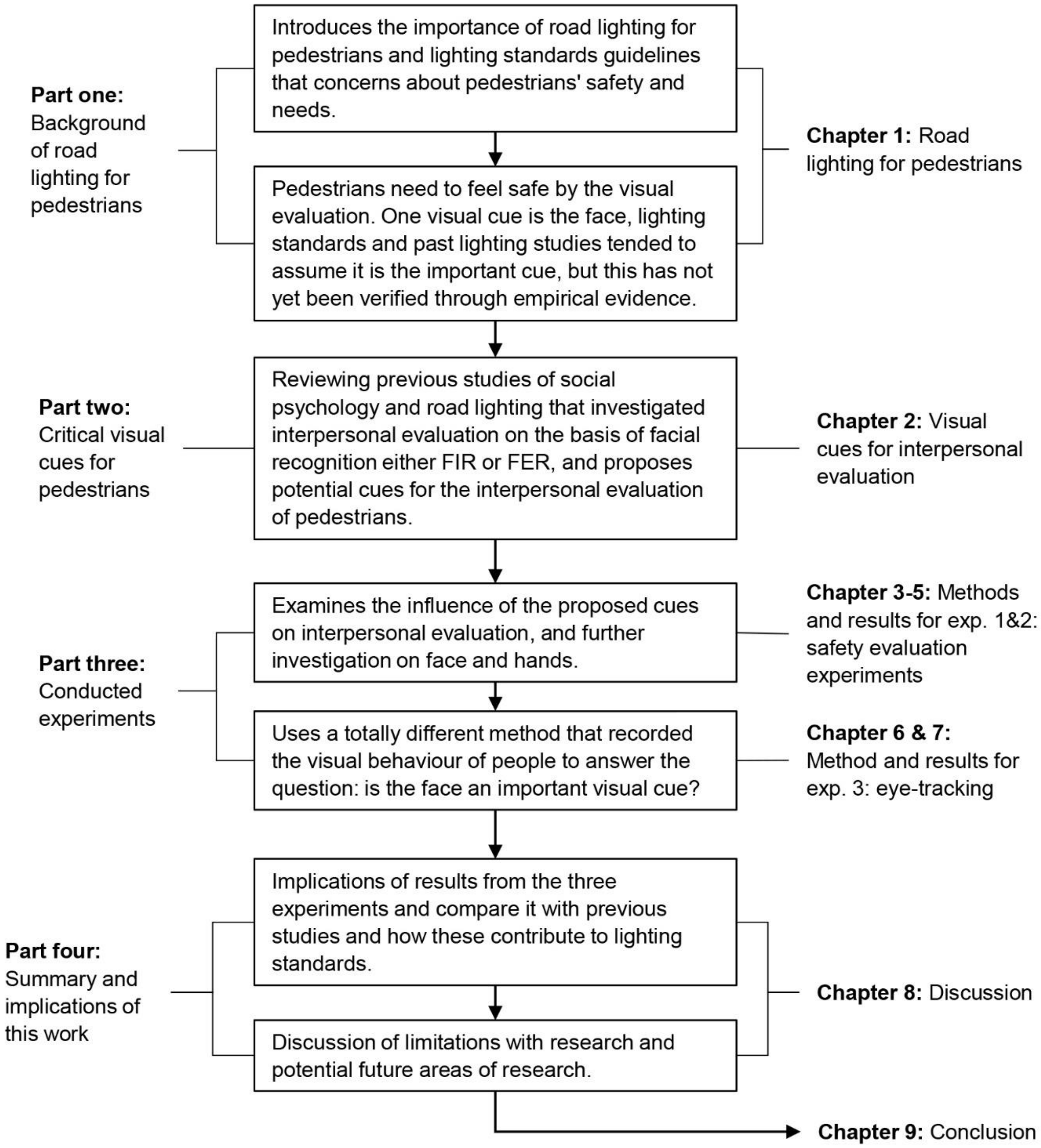


Figure 1.2. A summary of the thesis structure, and subject areas discussed in the chapters.

Chapter 2. Literature Review

2.1. Introduction

Chapter 1 explained the importance of road lighting in ensuring pedestrians' safety after-dark, and detailed current road lighting standards in the UK. It emphasised that the relationship between road lighting and interpersonal evaluation is not founded on robust scientific evidence (Fotios & Castleton, 2017; Fotios & Gibbons, 2018). The most important visual cue(s) used when evaluating other pedestrians after-dark are yet to be validated. This chapter reviews past studies (especially on lighting) and interpersonal evaluations.

2.2. Visual cues for interpersonal evaluations

This thesis focuses on visual cues about other people. After-dark, when visual performance is impaired, the ability to recognise people is reduced (see Section 1.2). Road lighting aims to enhance pedestrians' safety by assisting them in the process of interpersonal evaluation (BS 5489-1:2020) which is essential for guaranteeing their safety after-dark (Caminada & Van Bommel, 1984). It was also considered in one study to be the *"most important"* visual task undertaken by pedestrians; even if other pedestrians do not exhibit any sign of potential threat, they still need to be able to confidently make this evaluation (Simons et al., 1987).

There are several potential cues for pedestrians' interpersonal evaluations which could be investigated. Cook (1971, p. 65) discussed the information employed for these evaluations, and divided the visual cues into two groups: static and dynamic. Static cues do not change during encounters with other people, but dynamic cues do. Examples of static cues are the face, body, clothes, hairstyle, voice, make-up, and human adornments such as spectacles. Examples of dynamic cues are facial expression, body movement, gaze direction, orientation, distance, posture, gesture, tone of voice, and amount and fluency of speed. It is not known which, if any, of these cues plays the more dominant role.

Some of these cues have not been addressed in past studies because, typically, just one or two specific cue(s) have been investigated, and these have mostly been the face and/or body (Gifford, 1991; Knutson, 1996; Teufel et al., 2019). For instance, Willis et al.

(2011a, 2011b) investigated approachability according to the influence of facial expression, body posture, and gaze direction on approachability.

Attention in this thesis is first given to evaluations of the face, and those of the body, and the effect of gaze direction. This is because these cues have been used as observation targets, and are assumed to be relevant in past lighting studies. For instance, most lighting studies on interpersonal evaluation used only the face as a target (Alferdinck et al., 2010; Boyce & Rea, 1990; Caminada & Van Bommel, 1984; Dong et al., 2015; Fotios et al., 2017, 2018; Iwata et al., 2015; Johansson et al., 2001; Johansson & Rahm, 2015; Knight, 2010; Knight et al., 2007; Li & Yang, 2018; Lin & Fotios, 2015; Rahm & Johansson, 2018; Raynham & Saksvikronning, 2003; Rombauts et al., 1989; Romnée & Bodart, 2014; Yang & Fotios, 2015; Yao et al., 2009), whereas others used face and different cue(s); for example, Rea et al. (2009) investigated face and clothing; Okud & Satoh (2000) investigated face, cheek, and eye visibility; and Fotios, Yang & Cheal (2015) investigated face, body, and gaze direction.

In the studies which considered interpersonal evaluations broadly and did not address the impact of lighting, their observation targets are presented in Table 2.1. Some of these studies focused only on the face, some on the face and body, and others on gaze direction either alone or in conjunction with the face. Thus, the literature review commences with discussion of these cues: gaze direction, body features, and facial features.

Table 2.1. The observation targets in studies of interpersonal evaluations that do not address the impact of lighting.

Focus of observation	Studies
Face only	Johnson et al. (1991), Johnson & Mareschal (2001), Morton & Johnson (1991)
Face and body	Adams & Kleck (2003), Adams & Kleck (2005), Ekman (1965), Ekman & Friesen (1967), Van den Stock & de Gelder (2014), Willis et al. (2011a), Fletcher-Watson et al. (2008)
Face and gaze direction	Willis et al. (2011b), Strick et al. (2008)
Gaze direction only	Kendon & Cook (1969), Macrae et al. (2002)

2.2.1. Cues from gaze direction

Eye contact with other pedestrians has an influence on perceived safety, people usually feel uncomfortable and threatened when being stared at by strangers (Cook, 1977). When people look at the human face, they focus mostly on the eyes (Baron-Cohen et al., 1997, p. 39; Yarbus, 1997; Barton et al., 2001; Henderson et al., 2005; Walker-Smith et al., 1977); it is a special cue for social attention because people spontaneously look towards each other's eyes to reorient attention (Itier & Batty, 2009); and it is considered to provide more information (about complex mental states) than the mouth and to convey as much information as the whole face in evaluations (Baron-Cohen et al., 1997).

Gaze direction (otherwise known as eye movement or gaze behaviour) can send signals with respect to potential intentions; for example, at the beginning of a speech, the speaker makes initial eye contact as a welcoming signal, but then looks away as an indication of not wanting to be interrupted (Kendo, 1967); another example, people (including children) tend to infer that someone is thinking when his/her eyes are pointing away in an upwards direction (Baron-Cohen & Cross, 1992). Thus, people could be capable of understanding non-verbal language communicated by the eyes.

Two studies have found that direct gaze is regarded as a sign to approach, and averted gaze a sign to avoid (Ellsworth et al., 1972; Willis et al., 2011b). Studies conducted by Adams & Kleck (2003, 2005) and Willis et al. (2011b) found that gaze direction (whether direct or averted) is an important cue in social evaluation because it influences the perceptual processing of facial expressions of emotion. For example, angry expressions are identified more quickly when accompanied by direct gaze than by averted gaze. Similarly, faces combined with a direct gaze enhance gender identification faster than faces combined with averted gaze (Macrae et al., 2002), while attractive faces combined with a direct gaze enhance object evaluations more than when combined with an averted gaze (Strick et al., 2008).

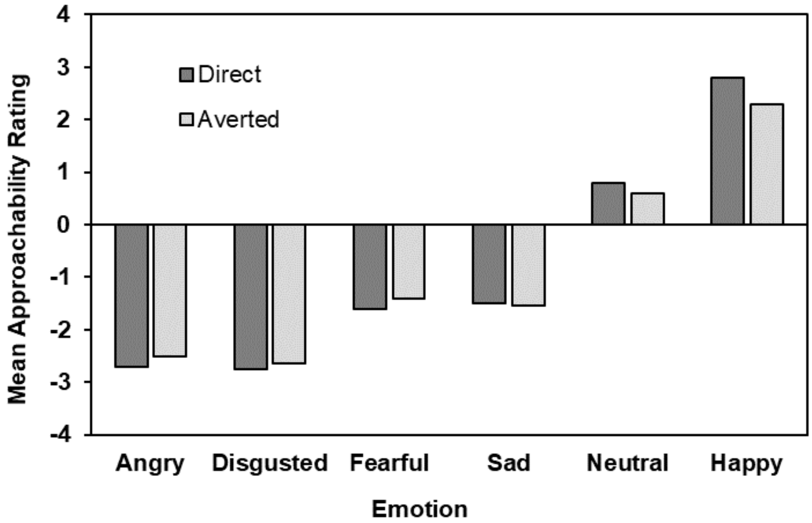
Adams & Kleck (2005) investigated how gaze direction influenced the perception of avoidance-oriented emotions (fear and sadness) conveyed through facial expressions. They conducted experiments asking the participants *"how likely the person depicted in the photograph is to experience each emotion".* Responses were given on a 7-point scale (1 = not at all frequently, to 7 = very frequently) to rate faces on four emotion scales (anger, fear, sadness, and joy). These were obtained from a range of reference facial images, and gaze direction in these images was digitally manipulated to be either direct or averted. The researchers found that gaze direction influences evaluations of emotional expressions, and that the prediction of emotion was enhanced by the direct gaze towards the eyes, especially when accompanied by specific facial expressions.

Willis et al. (2011b) used greyscale photographs of facial expressions portraying different emotions (angry, happy, neutral, disgusted, fearful, and sad) and displaying either a direct or averted gaze. The aim was to investigate how gaze direction effects approachability and evaluations (called trustworthy judgements). Participants were asked to perform an approachability task (asking for directions in a crowd), and a trustworthiness task (I would trust this person with my camera). The results revealed an influence of gaze directions that convey information for interpersonal evaluations; the perception of threat combined with emotional intensity played a critical part in interpersonal evaluations, and a direct eye gaze with angry faces was less approachable than an averted eye gaze with the same expression.

These studies that were undertaken by Adams & Kleck (2003, 2005) and Willis et al. (2011b) tested facial expressions and gaze directions as independent variables, but did not consider whether one or the other was more or less important. This can be estimated from the data presented in Figure 2 in Willis et al. (2011b), redrawn here as Figure 2.1 (a) which depicts the mean approachability responses. Participants were asked to rate a person's approachability by responding to the statement, *"I would approach this person to ask for directions"* on a 9-point scale (from -4 = strongly disagree to +4 = strongly agree). A significant effect of facial emotion was found ($p < 0.0005$), and happy faces were rated significantly higher (more approachable) than faces of all other expressions, whereas angry and disgusted faces were rated as less approachable (although the difference was not significant).

Figure 2.1 (a) depicts that the differences between different types of expression tend to be considerably greater than those between various gazes. For example, there is a difference on their scale of approximately 2.0 units between neutral and happy expressions, and a difference of 1.2 units between fearful and disgusted expressions. Figure 2.1 (b) depicts the difference between a direct and averted gaze for a neutral expression, which on their scale is approximately 0.2 units. This difference between direct and averted gaze was similar for all other facial expressions except for happy where the difference increased to 0.5. This might be interpreted as evidence that facial expression has a greater effect on interpersonal evaluations than gaze direction.

(a)



(b)

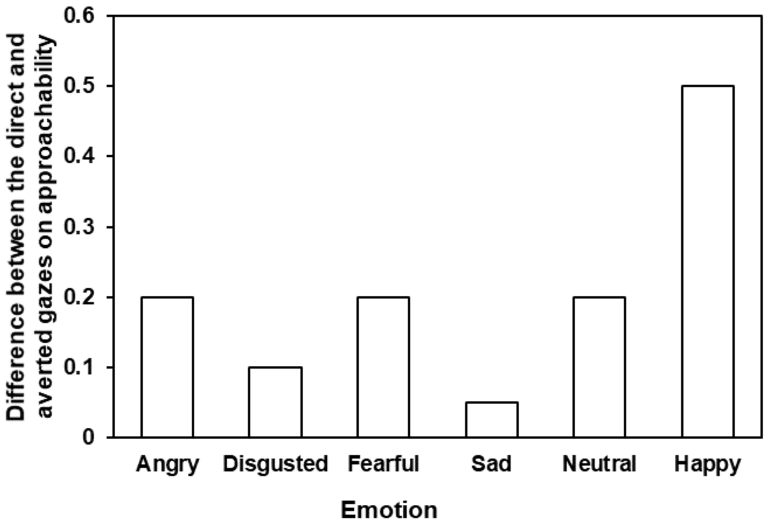


Figure 2.1. *(a)* Mean approachability ratings of emotional faces displaying direct and averted eye gaze; this is redrawn from Figure 2 in Willis et al. (2011b). *(b)* Estimated difference between the direct and averted gaze for each emotion portrayed by facial expression.

2.2.2. Cues from body features

Body language conveys an individual's attitude and expression of emotion. Two terms are used by scholars to describe this: body posture and body gestures. Body posture is defined as the position in which people hold their body while standing, walking or sitting, such as expanded, restricted, power, and submissive postures (Abele & Yzerbyt, 2021). Examples of different body postures are illustrated in Figure 2.2. Body gestures are defined as voluntary or involuntary movements of body parts such as head, arms, legs, including face, hands, and fingers (Kurien, 2010). These movements can indicate emotions such as anger, happiness, disgust, and fear (Thoma et al., 2013). Examples of different body gestures are depicted in Figure 2.3.

(a) **(b)**

Figure 2.2. Examples of body postures. *(a)* Male and female displaying power postures. *(b)* Male and female displaying a submissive posture. These images are Figure 3 in Abele & Yzerbyt (2021).

Figure 2.3. Examples of body gestures reflecting different emotions. This is part (b) of Figure 1 in Wu et al. (2021).

The distinction between posture and gesture is extremely small; normally, body posture is referred to as body gesture. As Bull (2016, p. 3) stated, body posture can be distinguished from body movements which are usually referred to as gestures. This causes inconsistency in the literature in that some studies refer to a body gesture as a body posture. For simplicity, this thesis uses the term body posture throughout.

Body posture can be regarded as a non-verbal form of communication (Remland, 2016). It is regarded as one of the most important elements in pedestrians' visual tasks (Simons et al., 1987), and might therefore be an important cue for interpersonal evaluation.

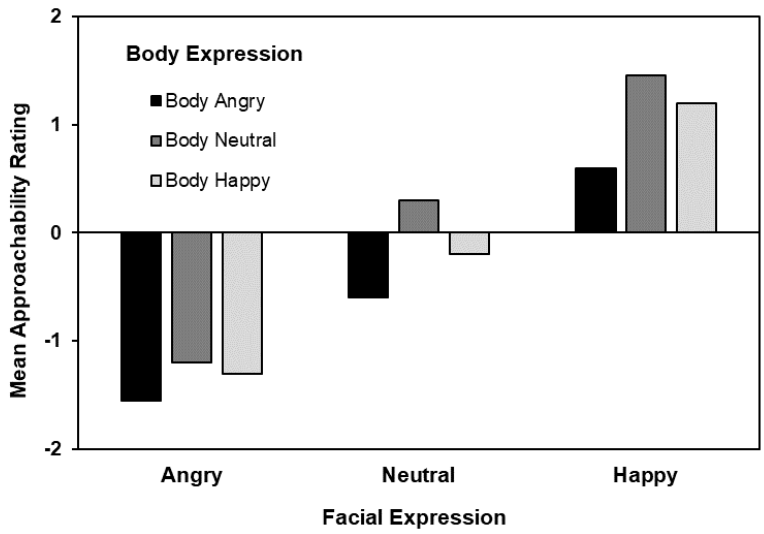
Willis et al. (2011a) demonstrated that body posture expressions (angry, happy, and neutral) influenced the decision as to whether to approach or avoid another person. The results suggested angry bodies were less approachable than the other expressions. The experiments not only investigated how approachability evaluations can be affected by bodies displaying an expression, but also by faces displaying the same three expressions. The experiments identified a significant main effect of angry faces as less approachable than happy faces, demonstrating that facial and body expressions play an important role in pedestrians' approachability evaluations.

In these experiments, participants were asked to perform two tasks. The first was an approachability task (asking someone for a direction on their way to meet a friend in an imagined crowded street situation) and then rating the approachability by responding to the statement, *"I would approach this person to ask for directions"* on a 5-point rating scale (-2 = strongly disagree, +2 = strongly agree). The second task was an emotion recognition task (indicate the body and the facial expression of the three emotions: angry, happy, and neutral) where responses were given by selecting the appropriate emotion displayed underneath each image.

The experiments used greyscale photographs of the bodies (and faces) of 10 actors (5 female, 5 male), all of whom displayed each of the three emotions. This meant 30 different bodies (and faces) were rated for approachability. At the beginning of each trial, a fixation cross was presented for five-seconds. The images were randomly shown one at a time at the centre of the screen against a white background. Along with the image, the statement and response scale were presented, and remained until a response was made by a mouse clicking, followed by a five-seconds interval period.

To investigate the impact of body expression, the face of the actor was digitally removed, with the body postures chosen to express the three different emotions. Conversely, to examine the effect of facial expression, the body was digitally removed, with facial expressions chosen to express the same three different emotions. However, the researchers did not state which cue had more influence on safety evaluation, whether body expression was less or more approachable than face expression. This can be estimated from Figure 2.4 (b) where, for a happy expression, the difference between face and body is approximately 1.2 units on their 5-point scale. This difference is close for the angry (0.8) and neutral expressions (0.4). The differences between various types of facial expression tend to be considerably greater than for body expression, as estimated from Figure 2.4 (a). For example, there is a difference of approximately 0.5 between neutral and happy body expressions and a 1.15 for neutral and happy facial expressions. This might be interpreted as evidence that facial expressions have a greater effect on interpersonal evaluations than body expressions.

(a)



(b)

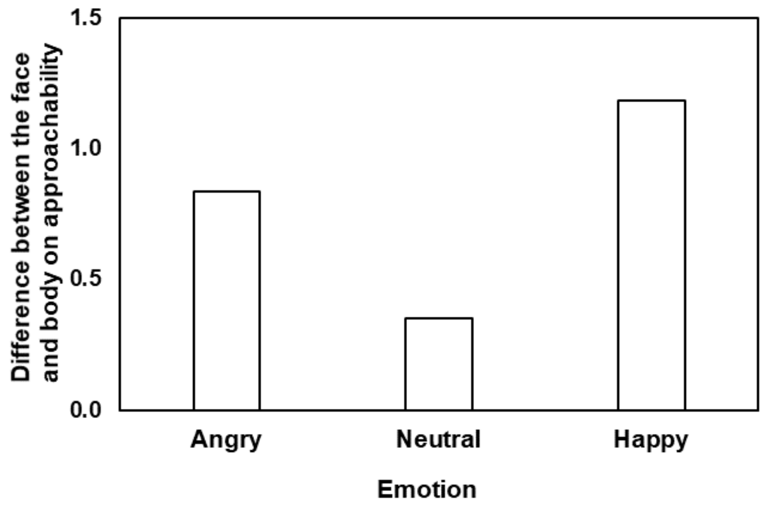


Figure 2.4. *(a)* Mean approachability ratings of emotional expression displayed by face and body; this is a redrawn graph based on Figure 5 in Willis et al. (2011a). *(b)* Estimated difference between the face and body for each emotional expression.

2.2.3. Cues from facial features

There is a tendency for people to look towards the face more than other elements of the person. This could be natural and innate behaviour driven by the immense face-processing skills that people possess from birth (Johnson et al., 1991; Johnson & Mareschal, 2001). Newborn infants prefer to look at faces more than other stimuli, which suggests that humans are born with some information about the structure of faces (Morton & Johnson, 1991). In addition, the human face conveys rich information about gender, age, emotions, and identity (Adolphs, 2002); and there is strong evidence that the face could be the primary cue for sending and showing emotions (Ekman & Friesen, 2003, p. 7). This section investigates the human face as a visible cue (more particularly, facial features), including the eyes which were discussed separately in Section 2.2.1.

Kirchner & Thorpe (2006) and Guyonneau et al. (2006), examined how fast the human eye can detect an animal using horizontal electrooculogram (EOG) electrodes to record eye movement. For stimuli, they used greyscale photographs of animals in their natural environments (mammals, birds, and fish) and distractor photographs (random landscapes without any animals such as forests, and man-made environments such as buildings) presented side-by-side in a dimly lit room. The images in their experiments were displayed for 20 and 30 ms, respectively. An example from Kirchner's & Thorpe's (2006) images is presented in Figure 2.5 (a). They employed a forced-choice saccade task that required participants to make a saccade as quickly as possible towards the side in which the photograph contained an animal. The saccade is defined as a rapid gaze location movement between two fixation points, and occurs when the eye gazes from one fixation to another (Holmqvist et al., 2011, p. 25). Their results revealed that a rapid visual performance was performed to extract an animal from its environment. Table 2.2 details the speed and accuracy of the detection responses.

Table 2.2. Results of the saccade-task of detecting an animal from photographs in two studies.

Study	Observation duration (ms)	Mean SRT* (ms)	Accuracy** (%)	Minimum*** SRT* (ms)
Kirchner & Thorpe (2006)	20	228	90.1	120
Guyonneau et al. (2006)	30	239.8	81.2	130

* SRT = Saccadic Reaction Time.
** Used statistical test (log-transformed percentage) of accuracy for the correct responses.
*** Saccade latency distribution was divided into 10 ms time bins. The minimum SRT reaction was the first bin, as this was where correct responses started to significantly outnumber erroneous saccades.

There is some logical sense to these findings. If people can rapidly detect an animal in scenes, they can easily identify other visual cues about a person on the road such as faces, bodies or holdable items. However, these studies used only animal targets, therefore, would the results be the same if the target were, for example, a person, a vehicle, or a piece of fruit?

Crouzet et al. (2010) used human faces instead of animals, conducting experiments using the same saccade-task implemented in the previous two studies (Kirchner & Thorpe, 2006; Guyonneau et al., 2006) except that the images were displayed for a longer period (400 ms). Two images were shown side by side in a dimly lit room, one with a target (showing faces, animals, or vehicles), and the other with a distractor (random neutral scene). An example from these images is presented in Figure 2.5 (b).

(a) **(b)**

Figure 2.5. Example of test image used in the saccade-task. *(a)* Test image used in Kirchner & Thorpe (2006); this image is part of Figure 1 in their paper. *(b)* Test image used in Crouzet et al. (2010); this image is part of Figure 2 in their paper.

The results of the saccade-task are depicted in Figure 2.6, which shows the distributions of Saccadic Reaction Time (SRT) between face, animal, and vehicle. This reveals a rapid processing of all the objects, with the face yielding the highest percentage of saccades with correct responses. The detection responses were quick and accurate for all the targets (Table 2.3). The face had the lowest mean SRT (147 ms) followed by the animal (170 ms), and the vehicle (188 ms), which means the face was recognised more quickly. It was also detected with a higher accuracy (94.5%) than the animal (82.3%) or the vehicle (75%). Nevertheless, these differences were not all statistically significant. For the mean SRT, the only significant difference was found between face and vehicle targets. For the minimum SRT, the face was lower (110 ms) than animal (120 ms) and vehicles (140 ms), which indicates that processing the face has an advantage over animals and vehicles when people need to discriminate it from distracting scenes.

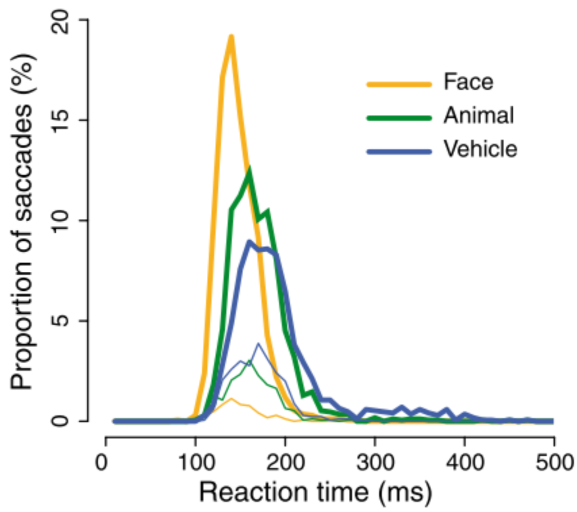


Figure 2.6. Results of the SRT distributions for the different targets: face, animal, vehicle. The thick lines indicate the correct responses, the thin lines depict the incorrect responses. This image is Figure 3 (top) from Crouzet et al. (2010).

Table 2.3. Results of the saccade-task of detecting different targets presented on images. The data in this table are generated from Figure 3 in Crouzet et al. (2010).

Target	Mean SRT* (ms)	Accuracy** (%)	Minimum*** SRT* (ms)
Face	147	94.5	110
Animal	170	82.3	120
Vehicle	188	75	140

* SRT = Saccadic Reaction Time.
** Used statistical test (log-transformed percentage) of accuracy for the correct responses.
*** Saccade latency distribution was divided into 10 ms time bins. The minimum SRT reaction was the first bin, as this was where correct responses started to significantly outnumber erroneous saccades.

Crouzet et al. (2010) also sought to determine whether the task can be reversed under voluntary control. For instance, can subjects switch between target categories, which means treating vehicles as targets and faces as distractors? The findings revealed that people tend to look towards faces, even if they were instructed to look at vehicles.

Therefore, these results indicate that the face is a more favourable visual cue than animals and vehicles. However, the experiments in these studies (Crouzet et al., 2010; Guyonneau et al., 2006; Kirchner & Thorpe, 2006) used greyscale photographs. Colour photographs display more information, and have a significant effect on people's memories (Kim, 2010).

Unlike Crouzet et al. (2010), Guyonneau et al. (2006), and Kirchner & Thorpe (2006), who used images of the person face only, and horizontal EOG electrodes for recording eye movements; Fletcher-Watson et al. (2008) used images of the whole body of the person (include face) and an eye-tracking device. Their work investigated the distribution of gaze in naturalistic scenes. Two images of scenes were presented side by side: one included a person and the other did not (random scenes from garden, office, living room, kitchen). The person was either sitting or standing, and their fixation was on a visible object or off-camera. An example from these images is presented in Figure 2.7.

Figure 2.7. Test image used in the study of Fletcher-Watson et al. (2008); this image is Figure 1 (a) taken from their paper.

Two test blocks were employed, free-viewing (FV) where no specific task was required (just simply looking), and gender-discrimination (GD) where participants were required to determine the gender of the person in the image. The results in Table 2.4 reveal that the percentage of total viewing and first fixation times were higher for the images containing a person than for the images without. This conclusion was the same for both blocks.

Table 2.4. Percentage of total viewing time and first fixations. These data are taken from Table 1 in Fletcher-Watson et al. (2008).

Percentage of	Blocks for images containing a person		Blocks for images without a person	
	Free-viewing	Gender-discrimination	Free-viewing	Gender-discrimination
Total viewing	59%	79%	36%	17%
First fixations	67%	82%	26%	13%

To provide greater insight into how fixations were distributed, four types of domains were analysed: face, body, background, and object (Table 2.5). The results of the FV block revealed a higher percentage of total viewing time and first fixations towards the face higher than the other domains. Similar conclusions were drawn for the GD block. The findings of Fletcher-Watson et al. (2008) suggest a strong tendency to gaze towards the person (particularly, directly to the face).

Table 2.5. Results for the time spent on each domain in images containing a person seen in two blocks. Percentage of total viewing time and first fixations. This is taken from Table 2 in Fletcher-Watson et al. (2008).

Test block	Percentage of the mean value for	Test domain			
		Background	Body	Face	Object
Free-Viewing (FV)	Total viewing	18%	12%	18%	10%
	First fixations	24%	27%	15%	1.0%
Gender-Discrimination (GD)	Total viewing	16%	23%	33%	8%
	First fixations	27%	26%	28%	1.2%

Other studies have investigated whether there is a strong relationship between face and body together in the interpersonal evaluation, where the movements of face and body, and their displayed emotions play important roles. For example, Ekman & Friesen (1967) presented evaluations of various emotions which with photographs of the subject's head only, body only, and head and body together. The results imply that facial expression identifies emotions, while body movements can indicate their intensity. In addition, it was reported that the interaction between facial expression and body posture in displaying emotions can be significant (Van den Stock & de Gelder, 2014).

A study by Willis et al. (2011a) revealed that the inferred meaning from body expressions is dependent on the equivalence with facial expressions. For example, angry bodies combined with angry faces can be identified more quickly than angry bodies combined with happy faces. Even though the experiments provided evidence to suggest that the face could be more important than the body in approachability, the researchers did not draw this conclusion.

To summarise, it has been established that the human visual perceptual process targets the face (for evaluations of either emotion or identity) which could result in it being the important cue for interpersonal evaluation. As noted previously, Cook (1971, p. 65) divided the visual cues into two groups: static and dynamic. When discussing the information used to assist personal evaluation, Cook argued that it is important to distinguish between ‘face’ and ‘face expression’. This aligns with Fotios & Johansson's (2019) suggestion that FER is a more effective way to evaluate the intentions of other people, they reviewed the basis of interpersonal evaluation and reported the following:

- There is some evidence to imply that the face is the most important cue with respect to intention.
- FER is a better method than FIR for evaluating the intentions of other people. One reason for this is that it removes the confound of familiarity.
- To date, there are no studies exploring the impact of lighting on recognition of hand gestures.

The facial identity task was reported to be easier than the facial expression task (Gao & Maurer, 2011). However, past lighting studies have focused on the face, operationalised as either FIR or FER, which will be discussed in the next section.

2.3. Past lighting studies of interpersonal evaluations

For pedestrians and road lighting, the visibility of cues is critical because it affects both their actual safety and their feeling of being safe. An eye-tracking study conducted to identify pedestrians' critical visual tasks suggested that interpersonal evaluations are an important consideration when setting standards for lighting, this is because pedestrians have a strong tendency to fixate on other pedestrians (Fotios, Uttley, et al., 2015). Therefore, in further research on lighting for interpersonal evaluations, it is important to establish the most relevant cues, as this knowledge will influence the measurements and how the task is operationalised, and thus affect the results and conclusions.

Figure 2.8 illustrates a situation where changes in lighting might affect a pedestrian's feeling of safety. The photographs depict a person ahead of the pedestrian, and how their location relative to the lamp-post (whether they are front-lit or back-lit) changes the visibility of the person (particularly the face). When they are front-lit, the face is visible, but when they are back-lit (the dominant source of light is behind them) their face is not visible, see Figure 2.8 (a) and (b), respectively. The lack of visual information about the back-lit person in this example means we are uncertain about them; hence, we may feel less confident about approaching that person. This is clearly an extreme example, but variations in illuminance on the face are also expected to change the ability to see facial details, and hence our ability to draw evaluations about a person.

Figure 2.8. Illustration of a pedestrian when front-lit and back-lit. It shows the effect of lighting on the appearance of a person (especially on facial features). *(a)* Front-lit person – most of the person is visible. *(b)* Back-lit person – most of the person is not visible.

This lack of visual information could be critical because the anonymity of the individual has an impact on perceptions of safety when evaluating an intended target (Schaller et al., 2003; Zhong et al., 2010). Similarly, a review of three studies of bank robberies involving interviews with incarcerated bank robbers and qualitative visual analysis based on recorded videotape demonstrated that invisibility is a desirable situation for criminals, as bank robbers like to select poorly lit places where they cannot easily be seen (Archea, 1985).

CIE report 136:2000 employed a similar lighting setting to Figure 2.8 to recommend lighting levels for urban areas that followed the lighting application fields of residential areas (Figure 2.9). This setting was then used to illustrate the impacts of light direction on visibility where the installation of road lighting in most cases was based on the appearance of people; front-lit and back-lit lighting (i.e. excessive or inadequate contrast) can distort this appearance. The report refers directly (and only) to faces without mentioning any other potential visual cues such as bodies or hands, a condition described as *"the effects of light direction on visibility and modelling of people's faces".* Note that in Figure 2.9 (b), the hand of the person is completely concealed which might affect the observer's evaluations as it makes it difficult to evaluate whether the target person is holding something of interest or a threatening object; hand visibility may therefore be a useful visual cue, since whether a person is holding an item or not influences evaluations of safety (Yang & Fotios, 2012).

Figure 2.9. Differences between front-lit and back-lit on the appearance of people. *(a)* Front-lit person – the face is visible. *(b)* Back-lit person – the face is not visible. This image is taken from the CIE report 136:2000, Figure 3.2, page 26.

Rombauts et al. (1989) also employed a similar lighting setting to that were shown in Figures 2.8 and 2.9 to investigate the relationship between semi-cylindrical illuminance and face recognition. In their experiment, participants were asked to walk towards a real target person and stop at certain distances to evaluate the visibility of the face. The evaluations were made using a 9-point rating scale (1 = not able to see anybody, 9 =

completely sure). Their results revealed that higher semi-cylindrical illuminance enhanced facial recognition ability at a distance of 4 m. Figure 2.10 illustrates a case where they described facial recognition as hardly feasible. Note that while they focus on the face, the hand of this person is concealed.

Figure 2.10. The lighting setting used to evaluate face visibility investigated by Rombauts et al. (1989), this image is Figure 3 from their paper.

To this date, the majority of previous research on lighting of interpersonal evaluations for pedestrians have tended to assume that the face is the important cue. Tables 2.6 and 2.7 summarise previous experiments investigating the effect of lighting on face-based evaluations. The majority of those studies have examined how changes in light level and/or Spectral Power Distribution (SPD) affect FIR. In those studies which examined light level, it was found that FIR increased at higher light levels (luminance or illuminance) (Alferdinck et al., 2010; Boyce & Rea, 1990; Caminada & Van Bommel, 1984; Dong et al., 2015; Okud & Satoh, 2000; Raynham & Saksvikronning, 2003; Rombauts et al., 1989). However, there are conflicting results for SPD in which some studies suggest a significant effect (Knight, 2010; Knight et al., 2007; Lin & Fotios, 2015; Raynham & Saksvikronning, 2003; Yao et al., 2009) and others do not (Alferdinck et al., 2010; Boyce & Rea, 1990; Rea et al., 2009; Romnée & Bodart, 2014).

Previous FIR studies in Table 2.6 have implemented one of three different procedures: identification, matching, and rating. The identification task involves detecting the identity of a person; for instance, to name a celebrity shown in a photograph. The matching task requires participants to match a target face with an array of reference faces. The rating task requires participants to assess the degree of recognisability of the target, using a scale which ranges from a low to a high degree of recognisability. This is usually performed at one or more fixed distances from the test participant to the target. An alternative is the stop-distance procedure where the participant is required to walk towards the target and stop when the assigned task can be completed; for example, when the participant is certain to the identity of the target.

An approximately equal number of those experiments have been conducted in the field and in the laboratory using observed targets as real person or photographs of faces. Combined with this is whether the target face was familiar to the participants, as the familiarity of the target can influence the process of identification (see Section 1.3). In terms of the time spent observing images, most studies used an unlimited duration, while others used a fixed set of observation times.

While the majority of lighting studies focus on FIR, a smaller sample of studies have examined how FER is affected by changes in lighting settings (Table 2.7). These experiments used an identification task to detect facial expression of emotions, where participants were required to identify the emotion conveyed by facial expressions (or to choose the correct emotion displayed in a photograph). This is usually conducted at one or more fixed distances (2, 4, 10, 15 m) from the participant to the target, or by using stop-distances. All these studies were undertaken in the laboratory using photographs displaying different facial emotions, whereas only one study used a 3D terracotta head of facial expressions. Regarding observation time, most of the studies used short durations (0.5, 1, 4 s), while others used unlimited time. The results of these studies (Fotios, Yang & Cheal, 2015; Fotios et al., 2017; Yang & Fotios, 2015; Fotios et al., 2018) revealed an effect of light level, but no effect of the SPD; in one study, the results were unclear (Romnée & Bodart, 2014).

Table 2.6. Past studies of lighting and face-based evaluations operationalised as FIR

Study	**Method**							**Effect found**	
	Carried out in	**Target**	**Familiar face***	**Distance between target and observer**	**Observation duration**	**Procedure**	**Task**	**Light levels**	**SPD**
Alferdinck et al. (2010)	Field	Real person	No	Stop-distance** (from 32 m)	Unlimited	Rating	Evaluate recognisability	Yes	No
Boyce & Rea (1990)	Field	Real Person	Yes	Stop-distance**	Unlimited	Facial recognition	Walking and observing	Yes	No
Lin & Fotios (2015)	Field	Photographs (celebrities)	Yes	Seven distances: 4 - 25 m	1, 3 s	Target identification	State name of person	Not tested	Yes
					Unlimited	Gender discrimination	State gender (male or female)		
Yao et al. (2009)	Field	Photographs (celebrities)	Yes	Stop-distance*** (from 25 m)	Unlimited	Identification of gender and name of celebrity	Say the name of person	Not tested	Yes
Rea et al. (2009)	Field	Real person	No	Stop-distance** (from 25 m)	Unlimited	Matching**	Walking and observing	Not tested	No
Knight et al. (2007)	Field	Photographs (celebrities)	Yes	Stop-distance** (from 15 m)	Unlimited	Facial recognition and colour identification	State names of person and colours	Not tested	Yes
Knight (2010)	Field	Photographs (celebrities)	Yes	Stop-distance** (from 15 m)	Unlimited	Gender discrimination and identity of the person	State gender and the name of person	Not tested	No

* Familiar face: yes: have seen the face before (friends, well known, or celebrities); no: not seen the face before.
** Stop-distance: participants were asked to walk towards the target or vice versa.
*** Matching: match the target faces with the reference face (pictures or screen).

TABLE CONTINUES ON THE NEXT PAGE

Table 2.6. (continued). Past studies of lighting and face-based evaluations operationalised as FIR

Study	Method							Effect found	
	Carried out in	Target	Familiar face*	Distance between target and observer	Observation duration	Procedure	Task	Light levels	SPD
Okud & Satoh (2000)	Lab	Real person	Yes	Seated at 1.5 m	Unlimited	Rating	Rate visibility of component	Yes	Not tested
Caminada & Van Bommel (1984)	Lab	Real person	Yes	Stop-distance**	Unlimited	Facial recognition	Identification (self-report)	Yes	Not tested
Dong et al. (2015)	Lab	Photograph (face sculptures)	No	Fixed at 10 m	Five durations: 0.1 - 10 s	Matching**	State name of person	Yes	Not tested
		Photograph (cel.)	Yes			Identification	Observation		
Rombauts et al. (1989)	Field	Real person	Not reported	Stop-distance*** (from 30 m)	Unlimited	Rating	Evaluate recognisability	Yes	Not tested
Romnée & Bodart (2014)	Field	Photographs (celebrities)	Yes	Stop-distance*** (from 36 m)	Unlimited	Identification and matching**	Walking and observing	No	No
Johansson et al. (2001)	Field	Real Person	No	Not reported	Unlimited	Walk in a footpath then answer a questionnaire	Can you recognize faces well?	Not tested	Not tested
Raynham & Saksvikronning (2003)	Lab	Real Person	Not reported	Stop-distance**	Unlimited	Facial recognition	Identification (self-report)	Yes	Yes
Iwata et al. (2015)*****	Lab	Real person	Not reported	Set distances: 3, 6, and 11 m	Unlimited	Facial recognition	Rating face visibility using answer sheets	Yes	Yes

* Familiar face: yes: have seen the face before (friends, well known, or celebrities); no: not seen the face before.
** Matching: match the target faces with the reference face (pictures or screen).
*** Stop-distance: participants were asked to walk towards a photograph.
**** No statistics reported on this study, there was a reported effect as a trend.

Table 2.7. Past studies of lighting and face-based evaluations operationalised as FER. All these experiments were conducted in a context of outdoor lighting (laboratory).

Study	**Method**							**Effect found**	
	Target	**Colour**	**Familiar face****	**Distance between target and observer**	**Observation duration**	**Procedure**	**Task**	**Light levels**	**SPD**
Fotios, Yang & Cheal (2015)	Photos*	Black & white	No	2, 4, 10 m	1 s	Observation of face & body	Forced-choice judgements	Yes	No
Yang & Fotios (2015)	Photos*	Coloured	No	4, 15 m	0.5, 1 s	Observation of face & body	Forced-choice judgements	Yes	No
Fotios et al. (2017)	Photos*	Coloured	No	4,15 m	0.5 s	Observation of face & body	Forced-choice judgements	Yes	No
Fotios et al. (2018)	Photos*	Coloured	No	4,15 m	0.5 s	Observation of face	Forced-choice judgements	Yes	No
Li & Yang (2018)	3D facial expressions	Brown-orange	No	4 m	4 s	Observation of face	Identify the expressions	Yes	No
Rahm & Johansson (2018)	Photos*	Black & white	No	Stop-distance*** (from 13.5m)	Unlimited	Observation of face	Stop walking when able to discern the facial expression	Yes	Yes
Johansson & Rahm (2015)	Photos*	Not reported	Not reported	Stop-distances ***	Unlimited	Facial emotion recognition and rating scale	Stop walking when able to discern the facial expression	Not clear ****	Not clear ****

* Photos: photographs of acted expressions.
** Familiar face: yes: have seen the face before (friends, well known, or celebrities); no: not seen the face before.
*** Stop-distance: participants were asked to walk towards a photograph.
**** Not clear: expression could be identified at a longer distance under the lighting application C. However, there was a varied lighting level and SPD which confounded this result.

A comparison of Tables 2.6 and 2.7 reveals conflicts in the results. Some FIR studies suggest that changes in SPD affect the evaluation and some do not, while none of the FER studies reveals such an effect. The differences between FIR studies are presented in Table 2.8. All the studies reporting an SPD effect used photographs of familiar faces of celebrities as a target (two studies did not report the familiarity), while the studies that did not find an effect used familiar and unfamiliar target faces with differences between a real person and photographs. This is, however, not enough evidence to confirm that familiarity with target faces affects SPD.

Table 2.8. Studies of lighting and FIR and whether an effect of SPD was found.

Target	Familiar or unfamiliar face?	Effect of lamp SPD	Studies
Real Person	Not reported	Yes	Raynham & Saksvikronning (2003), Iwata et al. (2015)
Photographs (celebrities)	Assumed familiar	Yes	Lin & Fotios (2015), Knight et al. (2007), Yao et al., 2009
Photographs (celebrities)	Assumed familiar	No	Knight (2010), Romnée & Bodart, 2014
Real person	Familiar	No	Boyce & Rea (1990)
Real Person	Unfamiliar	No	Alferdinck et al. (2010), Rea et al. (2009)

To summarise, interpersonal evaluations for pedestrians' safety are based on the visual cues of other people, such as the face, body, hands, and gaze direction. However, it is unknown (or has not been confirmed) which cue is the most important. To be certain about the important visual cue is critical for road lighting installation because pedestrians' safety (more specifically, the feeling of being safe) may be affected differently depending on the lighting condition falling on that cue(s).

If FIR is not an appropriate operation for lighting research (see Section 1.3), this may lead to incorrect recommendations in lighting guidance. In addition, the discussion regarding the lack of visual information in a back-lit condition (Figures 2.8 to 2.10) raised a problem that only the face was investigated, whereas hands could also display a threat, as well as gaze direction and body movement (see Sections 2.2.1 and 2.2.3 respectively). This means that the choice of visual task is an important question for experimental design. Therefore, this thesis examines whether the human face is the truly important visual cue for safety evaluation.

A review of the literature reveals important limitations in previous studies; first, the face was thought to be the most important cue in many studies (e.g. Fletcher-Watson et al., 2008; Johnson et al., 1991; Johnson & Mareschal, 2001), but none of these studies had validated this idea. Second, some studies (e.g. Guyonneau et al., 2006; Kirchner & Thorpe, 2006) forced the participants to look at the face (or other stimuli, e.g., animals) where attention was involuntarily biased towards faces. For instance, the image of the person (specifically the face) overwhelmed the scene by (for example) being in a large size. Third, two studies conducted by Willis et al. (2011a, 2011b) have supplied evidence that facial expression and body posture are important cues for interpersonal evaluation, but these were limited to testing only one factor. Finally, no study has examined perception of safety related to interpersonal evaluations in a road lighting environment after-dark. It is assumed that being better able to make visual evaluations about other people leads to a higher level of safety, but this is yet to be validated.

2.4. Suggested cues for interpersonal evaluations

The literature (particularly, experiments involving pedestrian lighting) does not appear to provide authoritative guidance as to which visual cue(s) is/are the most important for interpersonal evaluations between pedestrians. As an alternative source of evidence, we can consider the set of personal awareness strategies initially suggested by Nancy Clanton which were raised during discussion within the IESNA outdoor lighting committee. Clanton is a senior lighting designer from the USA, and the CEO of Clanton & Associates, a lighting design firm she founded in 1981. From this personal experience in life and lighting, Clanton proposed a sequence of cues. She stated that (Fotios et al., 2019): *'I personally check*

- *Gender (women, or man/woman together gives me the least anxiety)*
- *Then number (2 or 3 against 1 is high anxiety)*
- *Eye contact (are they looking at me and scoping me out?)*
- *Walking direction (are they walking towards me to block my path?)'.*

Although these cues were advocated by an expert lighting designer, it remains a personal opinion and needs to be tested against the literature. Accordingly, the effect of the gender of the evaluated person does not appear to have been addressed in previous work, that is whether we feel more fearful being approached by a male or by a female. By contrast, the effect of the gender of the evaluator has been addressed in numerous studies (Boomsma & Steg, 2014; Fisher & Sloan, 2003; Foster et al., 2004; Gover et al., 2011; Koskela, 1997; Lebowitz, 1975; Lee, 1982). These studies reported that, walking outdoors after-dark and in a low lighting condition, females are less likely to feel safe than males.

Regarding the number of people, very few studies have investigated the effect of this has on perceived safety. For example, if an individual walks alone after-dark, a potential encounter with one person would be safer than encountering more than one. One study by Nasar & Jones (1997) asked participants to comment on aspects of the surrounding environment that made them feel safe or unsafe. The statistical analyses were related to fear and safe spots, but no one mentioned the impact of a larger number of people in relation to fear. Nevertheless, walking in the company of a friend provides more safeguards and increases the feeling of being safe (Staats & Hartig, 2004), hence, a group of people is more willing to take risks (and face danger) than an individual by him/her self. However, no study has confirmed whether one person or more than one gives the observer a safer feeling.

Eye contact has been discussed as an important cue in interpersonal evaluation (see Section 2.2.1). The literature review on eye movements confirms that gaze direction (direct or averted) has an effect on social judgement, and a direct gaze is more likely to give the observer a sense that a person is safe to approach than an averted gaze (Willis et al., 2011b). Nevertheless, the eye contact as a visual cue was excluded from the investigation of this thesis because it is an assessment of gaze direction towards rather than away from the observer, and it is difficult to evaluate as a visual cue because the pupil is a very small target, especially after-dark when vision becomes more difficult.

Walking direction and safety is rarely addressed in the literature as most studies (Chandra & Bharti, 2013; Dias et al., 2015; Fotios, Yang, & Uttley, 2015; Knoblauch et al., 1996) discussed the distance between people or the walking speed, rather than whether they are walking towards the observer or away from them. Thus, based on the review of the literature and the recommended personal awareness strategies by the lighting designer (Clanton), the following visual cues are recommended to evaluate their effect on interpersonal evaluations of pedestrians:

- Gender
- Number of people
- Walking direction
- Light direction: fall of light direction on the persons.
- Exposure of face and hands: how well they can be seen.

2.5. Research hypotheses

The aim of this study was to identify the important visual cue(s) used by pedestrians when making interpersonal evaluations about other pedestrians when walking after-dark; this supports the application of road lighting in minor roads that primarily targets the needs of pedestrians (BS 5489-1:2020). This is a worthwhile objective for providing accurate lighting standards and appropriate road lighting that will enable safe walking. To meet this aim, and after the literature review, the following hypotheses were addressed:

H1: The face as a visual cue has an influence on interpersonal evaluations.

H2: The face is a more important visual cue in interpersonal evaluations than the body or the hands.

These hypotheses were tested by conducting three experiments, the structure of which is illustrated in Figure 2.11.

Past studies of interpersonal evaluation have assumed that the face is the most critical cue, and testing that focus is the aim of the current work. However, this is not assumed to be a universal conclusion. In some cultures, such as Japan, the Philippines, and Vietnam, looking towards the faces of others is not common and considered unacceptable. For example, in Japanese culture, children are taught in school to not look at the face and instead look at the neck (Morsbach, 1973); staring at the face is considered rude in the Philippines, and it is considered suspicious or threatening in Vietnam (Kawaguchi-Suzuki et al., 2019).

If the face is not the important cue then some other object or body feature must be instead. Further work should therefore establish what those cues are when it is not the face. Other features may subtend a different size at the observer’s eyes than does the face, and objects of different size exhibit a different relationship between light level and visual performance (Boyce, 2014, p. 136, Figure 4.12), and this may lead to a different interpolation of the optimal lighting conditions for the task.

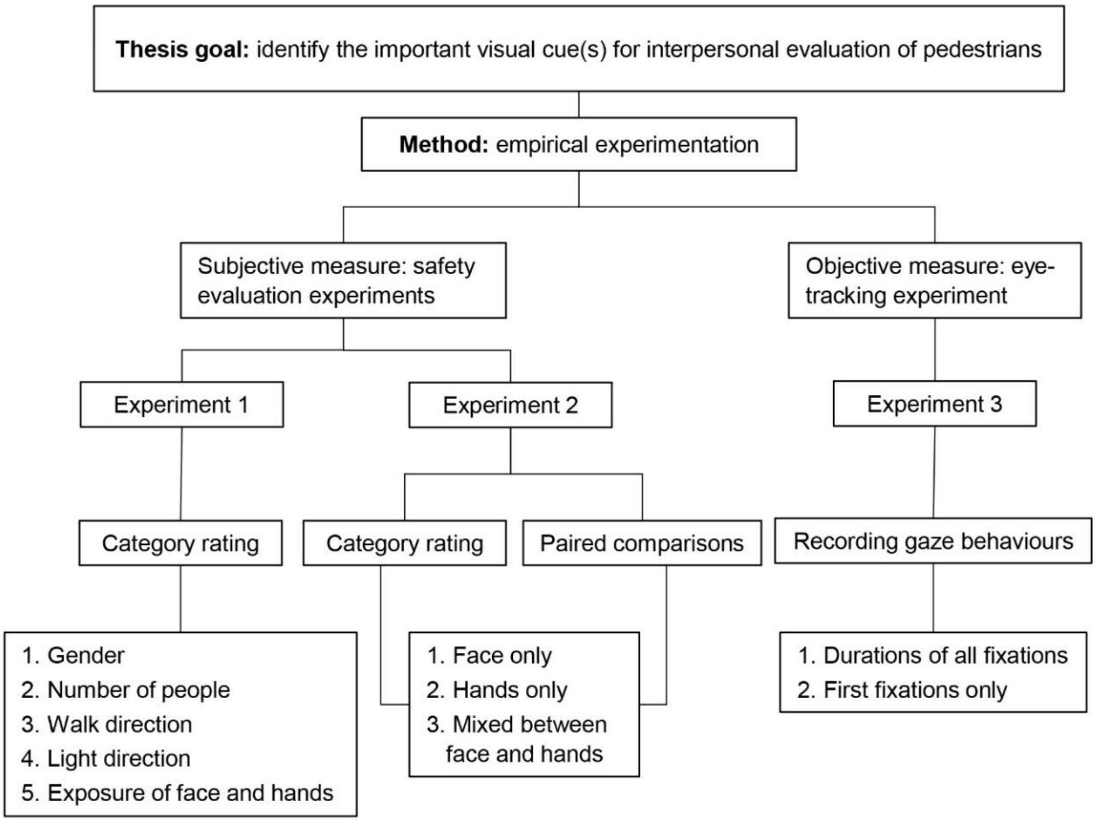


Figure 2.11. Summary of the methodological structure employed in this thesis.

Chapter 3. Safety Evaluation – Experiments 1 and 2: Method

3.1. Introduction

The literature review revealed that although the face has been used as a target in multiple lighting experiments, there is insufficient evidence to confirm that it is the primary focus of attention when evaluating the intentions of other people.

This chapter describes the method of Experiments 1 and 2 that carried out to investigate the key visual cues used by others when performing interpersonal evaluations of pedestrians. A subjective measure was employed whereby participants were asked to rate their feelings of safety regarding an approaching person(s), and to select a safer situation from a pair of images. The purpose of Experiments 1 and 2 was to investigate the research hypotheses:

H1: The face as a visual cue has an influence on interpersonal evaluations.

H2: The face is a more important visual cue in interpersonal evaluations than the body or the hands.

3.2. Apparatus

The perceived interpersonal safety of pedestrians was investigated by testing a range of potential visual cues. These were portrayed in a series of images that were evaluated by participants. The images were generated by capturing photographs of outdoor scenes after-dark, into which were embedded photographs of actors portraying each cue.

The devices used in the photography sessions were a Canon D70 camera (Experiment 1) and an iPhone X camera (Experiment 2). Although the image resolution of the iPhone X camera is lower than that of the Canon D70, it nevertheless produced images with a resolution exceeding that required for the experiment. The final sizes of images presented in both experiments were smaller than the photographs taken with both devices (Table 3.1).

Table 3.1. Details of the camera devices used in the safety evaluations – Experiments 1 and 2.

Camera device	**Used in**	**pixel size***	**Size (W x H) of the photograph taken**	**Size (W x H) of images presented in the two procedures**	
				Rating	**Paired**
Canon D70	Experiments 1	20.2 MP*	5472 × 3648 PPI**	1559 x 875 PPI**	Note used
iPhone X	Experiments 2	12 MP*	3024 × 4032 PPI**	1559 x 875 PPI**	

* MP = Megapixels
** PPI = Pixels Per Inch

After-dark scenes of several locations in Sheffield UK were photographed to enable subsequent determination of which scene(s) to use (for examples see Appendix A). The final scene chosen met the following criteria: a subsidiary road; a clear and flat pedestrian path; no apparent glare in the near field; and no other people nearby, or other people at a sufficiently distant to be easily removed by image editing. These criteria are important because a subsidiary road is where people can meet pedestrians, a flat pedestrian path is important for reducing complexity, and because glare affects visibility.

The background scene used in Experiment 1 is depicted in Figure 3.1. In this experiment, differences in evaluation due to the effect of the scenes (street appearance) were not assessed, and therefore only one scene was used.

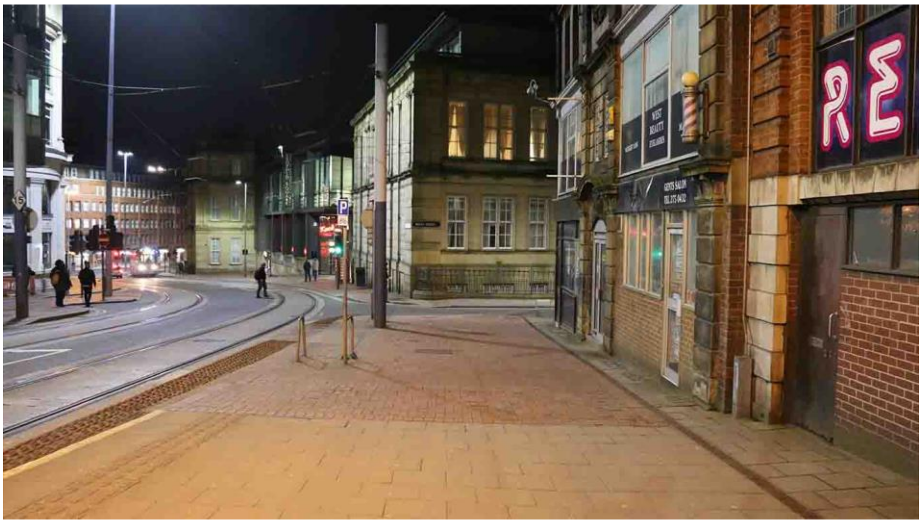


Figure 3.1. The background scene used in the test images of Experiment 1. There are no pedestrians nearby, and no digital manipulation was performed on this image.

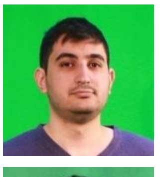
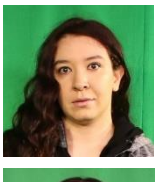
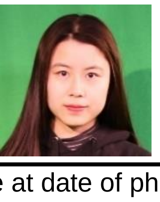
Experiment 2 used two background scenes in test images to determine whether these influenced the safety evaluations; where the type of the environment is one factor that could affect pedestrians' feeling of safety, for example, lighting was considered to be the most important environmental feature in making a place feel safer (Loewen et al., 1993), and a lower lighting and/or an entrapment settings make people feel less safe (Boomsma & Steg, 2014; Fisher & Nasar, 1992; Nasar et al., 1993; Nasar & Jones, 1997). The two background scenes were photographed at locations in Sheffield. Figure 3.2 (a) presents the first scene, which was a brightly lit road, dominated by HPS streetlamps which had previously been used as the background for Experiment 1; this background was adjusted by removing a few pedestrians and a tram (an error found in Experiment 1). Figure 3.2 (b) presents the second scene that was chosen to have different light environment than first scene, so they can be compared; this scene was a dimly lit back alley dominated by LEDs.

Figure 3.2. The two scenes used as backgrounds in Experiment 2: *(a)* Scene 1 shows a lit road with no pedestrians. *(b)* Scene 2 shows a dimly lit road with no pedestrians.

Postgraduate students were used as actors. Using multiple (four or five) actors rather than only one provided more images for comparison, and increasing the generalisability of the findings. All participants agreed to be involved in this study, and each signed a written consent form allowing the experimenter to use their images for research purposes in the thesis and for publication (see Appendix F). The images of the actors were subsequent superimposed against the background scene(s).

For Experiment 1, photographs of the actors were captured in a media studio at the University of Sheffield. The studio contains a green background curtain to enable easy image cropping and diffuse lighting conditions in order to avoid strong shadows being cast on the actors. Five actors were involved, they comprised two males and three females ranging in age from 20 to 33 years; their nationalities were Saudi (males 1,2), Chinese (Male 3, Female 2), and Turkish (Female 1) (Table 3.2).

Table 3.2. Information about the actors used to create the test images of Experiment 1 compared with the typical body size of the British citizen.

Actor reference	Image of their face	Home nationality	Age (y)*	Height (m)*	Weight (kg)*	Typical height (m)**	Typical weight (kg)**
Male 1		Saudi Arabia	33	1.72	72	1.78	79
Male 2		Saudi Arabia	33	1.78	98	1.78	79
Male 3		China	26	1.70	52	1.78	79
Female 1		Turkey	31	1.75	76	1.64	70
Female 2		China	24	1.60	47	1.64	70

* Approximate at date of photograph
*** Source: www.onaverage.co.uk

For Experiment 2, photographs of the actors were taken in an office with diffused artificial and natural lighting. Four actors were involved, they comprised two males and two females ranging in age from 20 to 30 years; their nationalities were Jordanian (Male 1), British (Male 2), Turkish (Female 1), and Portuguese (Female 2). For these photographs, the actors wore two types of clothing: a blue jacket and a yellow hood, and their body size was within the average range for British citizens (Table 3.3).

Table 3.3. Information about the actors used to create the test images of Experiment 2 compared with the typical body size of the British citizen.

Actor reference	Image of their face	Top clothing*	Home nationality	Age (y)**	Height (m)**	Weight (kg)**	Typical height (m)***	Typical weight (kg)***
Male 1		Blue jacket	Jordanian	29	1.71	78	1.78	79
Male 2		Yellow hood	British	26	1.72	67	1.78	79
Female 1		Blue jacket	Turkish	28	1.74	68	1.64	70
Female 2		Yellow hood	Portuguese	30	1.70	66	1.64	70

* All actors wore blue jeans
** Approximate at date of photograph
*** Source: www.onaverage.co.uk

To resemble the tested cues, different poses were taken in a simulated walking position where the right leg was forward and the left leg behind (or the reverse). Each pose required the actors to face straight forward, hands by their side, with no watch or other accessories worn, and a neutral facial expression. Within this position, they were asked to portray different poses: Experiment 1, facing towards or away from the camera and to wear/not wear a face covering (jacket with a hood) and hands at their sides or placed in pockets (Figure 3.3); Experiment 2, displaying a range of face and/or hand exposures (Figure 3.4).

Figure 3.3. Example of the different poses of the actors photographed in the media studio for Experiment 1. This shows *from the left*: Female 1 walking towards the camera, her face and hands partially concealed; Female 2 walking away from the camera, her face and hands uncovered; Male 1 walking towards the camera, his face and hands exposed; Male 2 walking towards the camera, his face and hands partially concealed; Male 3 walking away from the camera, his face and hands concealed.

Figure 3.4. Example of different poses of the actors photographed in an office for Experiment 2. Shows *from the left*: Male 1, face and hands exposed; Male 2, lower part of face concealed and hands exposed; Female 1, face exposed and hands concealed; Male 1, top part of face concealed and hands exposed; Female 2, face and hands concealed.

Adobe Photoshop software, which facilitates the digital manipulation of images, was used to embed the images of actors into the background scenes. All images required editing (cropping, resizing, adjusting brightness, creating actor's shadows, improving colour, and adding desired effects) to finally merge and fit them into the scene photographs. Figure 3.5 depicts one test image before and after manipulation.

Figure 3.5. Example of image manipulation. *(a)* The original image shows Female 1 in the media studio before editing. *(b)* The test image (no. 3) shows Female 1 embedded into the scene after cropping the green background and adding her shadows on the ground.

The method of embedding the actor’s image on the scene to be observed at that distance was achieved by taking a real image of a target at the required distance, then, using Photoshop, embedding the actor’s image into the scene at the exact measured distance from the target and the observer (Figure 3.6). The same principle was used to determine the size of the shadows taken from the real photograph. This meant the actors’ shadows were inserted realistically into the surroundings of the scene (Figure 3.6).

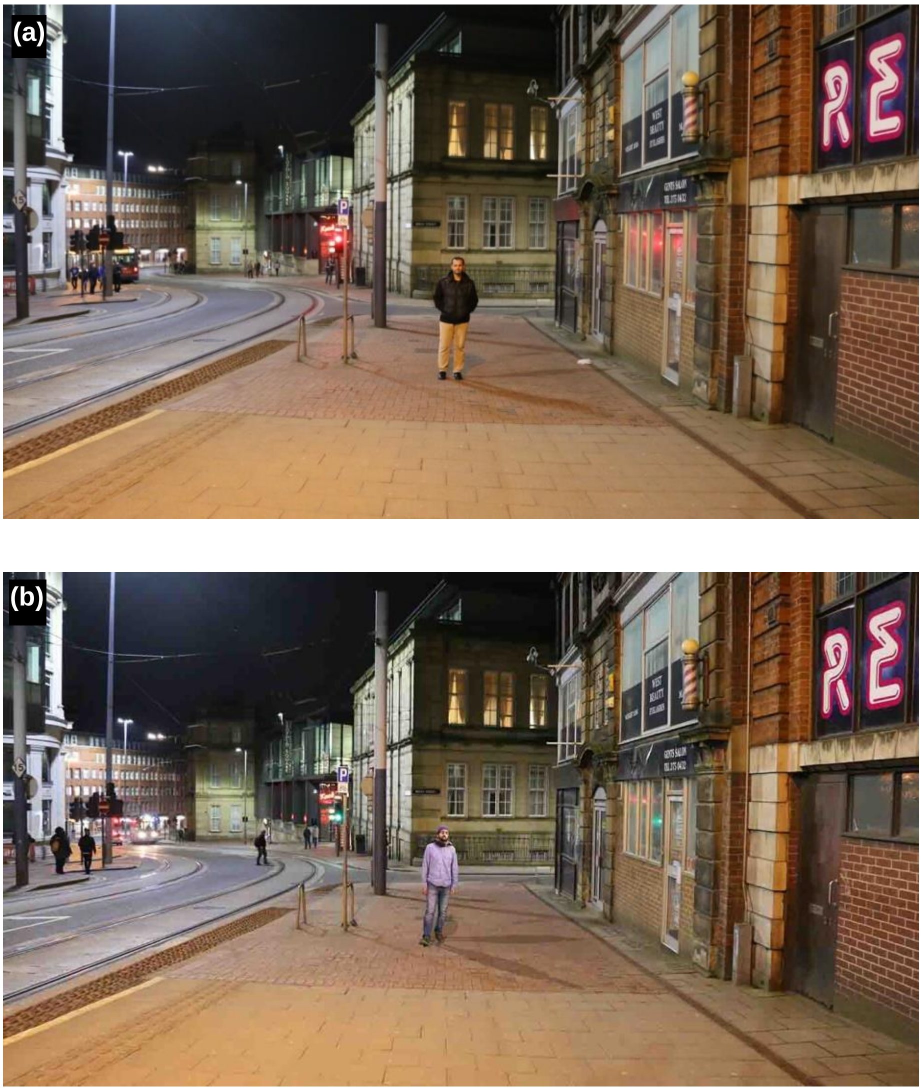


Figure 3.6. Embedding the target at 9 m distance and creating the actors' shadows. *(a)* The real photograph without any manipulation for a person 9 m away from the camera. *(b)* The test image (no. 1) used after embedding the actor’s photograph taken in the media studio into the photograph of the background scene. The distance and shadows were created based on the real photograph of Image (a).

The lighting effect on the actors gives the impression of being either front-lit or back-lit; this effect was achieved by dimming the brightness of the actor to display the back-lit condition or leaving the target without any dimming to display the front-lit condition (Figure 3.7).

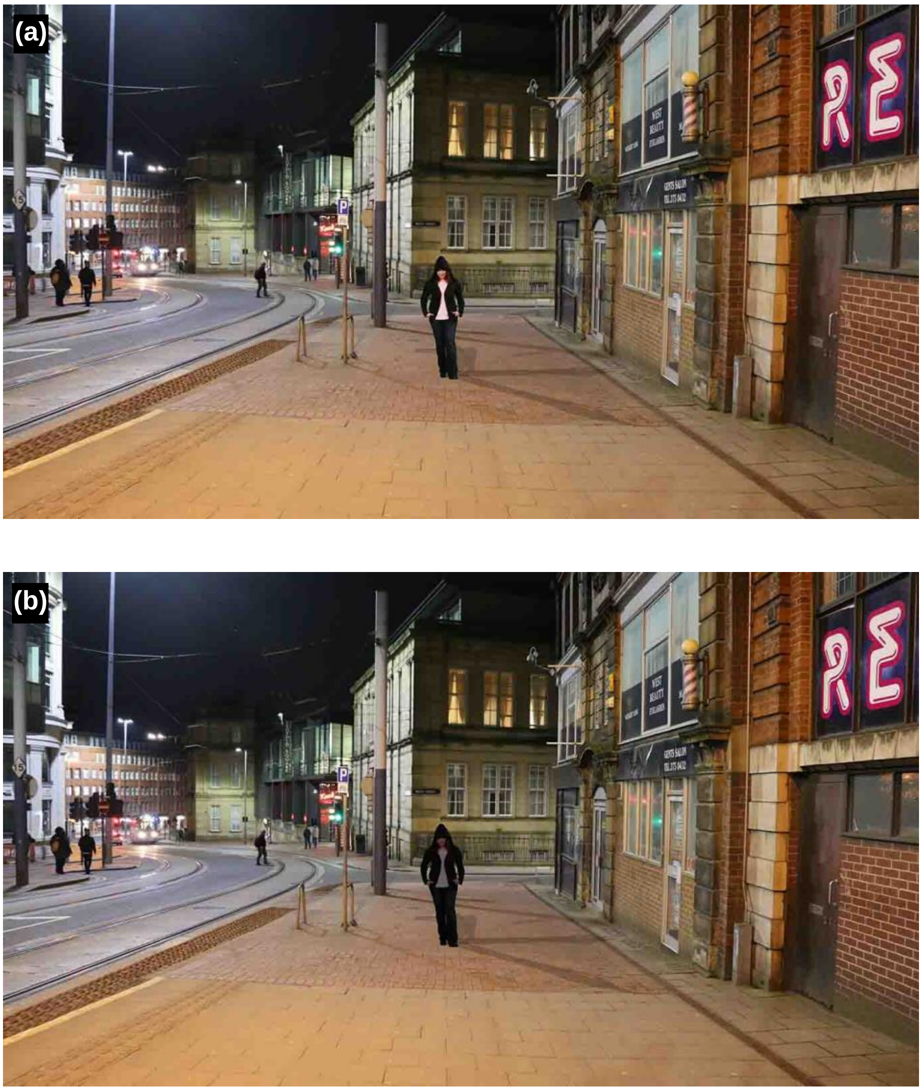


Figure 3.7. Creating the impression of the back-lit condition using Photoshop software. *(a)* The test image (no. 4) used shows the target in a front-lit condition where the actor is bright without any dimming. *(b)* The test image (no. 8) used shows the same target but in a back-lit condition where the visibility of the target was reduced by dimming the brightness on the actor.

The actors' position and size were adjusted to simulate a distance between a real person and the observer of approximately 9 m and 4 m (Experiments 1 and 2, respectively). The decision to change the observed distance in Experiment 2 was made after a larger distance in Experiment 1 was found to make face and hands less easy to distinguish (Figure 3.8). In addition, the distance of 4 m has been recommended as the minimum distance between pedestrians to enable facial recognition and expressions (Johansson & Rahm, 2015; Yang & Fotios, 2012), and for them to take any necessary defensive action such as walking away from a person or applying evasive measures if threatened (Caminada & Van Bommel, 1984; CIE, 2000; Hall, 1966, p. 123; Rombauts et al., 1989).

Figure 3.8. Illustration of different distances of the target. From the left: 15 m, 12 m, 9 m, 6 m, 4 m, and 3 m. Experiment 1 used a distance of 9 m, and Experiment 2 used a distance of 4 m.

Figure 3.9 (a) depicts the interpersonal distance represented at 4 m in a real street, observation of a target height of 1.7 m (as an average height of a person, source: www.onaverage.co.uk) at this distance would subtend an angle at the observer's eye of 24°. Figure 3.9 (b) depicts the interpersonal distance in the laboratory, the observation distance from the eye to the wall was 1.5 m. Thus, the required target size on the wall was 0.64 m with the same subtended angle of 24°.

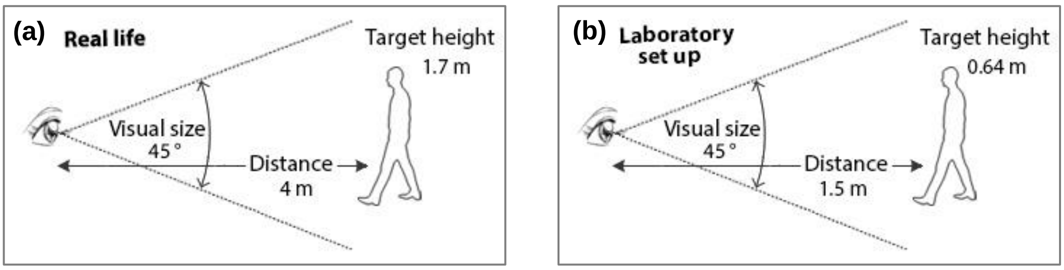


Figure 3.9. Calculation of the interpersonal distance. *(a)* Dimensions of the real size in the real life. *(b)* Dimensions of the projected target size in the laboratory.

The test images were presented in Experiment 1 using a 27-inch PC monitor (iiyama ProLite E2409HDS) which displays a resolution of 1920 x 1080 pixels per inch (PPI). Images of Experiment 2 were presented using a projector (Optoma HD31UST) to cast the images onto a white wall. This projector provides a high-quality image with a resolution of 1920 x 1080 pixels and dimensions of 2.21 m width and 1.25 m height (approximately 1.15 mm per pixel, equivalent to a 131.4-inch monitor). The projector was placed above a table 0.75 m above the ground. The distance between the projector and wall was 0.25 m; this short distance is possible due to the ultra-short-throw lens. The final size of images presented via the PC monitor and the projector were smaller in resolution, but different in their dimensions (Table 3.4).

Table 3.4. Details of the display devices used in the safety evaluation – Experiments 1 and 2.

<table>
<tr><th rowspan="3">Experiment</th><th rowspan="3">Display device</th><th rowspan="3">Maximum resolution size of the device</th><th colspan="4">Size of presented images (W x H)</th></tr>
<tr><th colspan="2">Pixels Per Inch</th><th colspan="2">Meter</th></tr>
<tr><th>Rating</th><th>Paired</th><th>Rating</th><th>Paired</th></tr>
<tr><td>One</td><td>PC monitor 27-inch</td><td rowspan="2">1920 x 1080 PPI*</td><td rowspan="2">1559 x 875 PPI*</td><td>Not used</td><td>0.60 x 0.34 m</td><td>Note used</td></tr>
<tr><td>Two</td><td>Projector</td><td>935 x 525 PPI*</td><td>2.21 x 1.25 m</td><td>2.21 x 1.25 m</td></tr>
</table>

* PPI = Pixels Per Inch.

3.3. Stimuli

Experiment 1 was conducted to compare the relative importance of different visual cues proposed in the literature (see Section 2.7). These are important, if available, in establishing a hierarchy of safety factors for pedestrians. The five potential cues and the predicted safer condition for each are described in Table 3.5.

Table 3.5. Proposed visual cues and the predicted condition leading to feeling safer.

Visual cues (of the approaching person)	Predicted safer condition for the approaching person
Gender	Female rather than male
Number of people	One person walking alone rather than two together
Walking direction	Walking away from the observer rather than towards him/her
Light direction	Front-lit rather than back-lit condition
Exposure of face & hands	Exposed rather than concealed face and hands

A combination of test cues and actors was established. This gave a total of 16 test images captured with different variations (Table 3.6). Example images are presented in Figure 3.10, and the entire set can be found in Appendix C, Figure C.1.

Table 3.6. Variations in the 16 test images used in Experiment 1.

Image no.*	Gender	Number of people	Walking direction	Light direction	Exposure of face and hands
1	Male 1	One person	Towards	Front	Exposed
2	Male 2	One person	Towards	Front	Partially concealed
3	Female 1	One person	Towards	Front	Exposed
4	Female 1	One person	Towards	Front	Partially concealed
5	Male 1	One person	Towards	Back	Exposed
6	Male 2	One person	Towards	Back	Partially concealed
7	Female 1	One person	Towards	Back	Exposed
8	Female 1	One person	Towards	Back	Partially concealed
9	Male 3	One person	Away	Front	Exposed
10	Male 3	One person	Away	Front	Partially concealed
11	Female 2	One person	Away	Front	Exposed
12	Female 1	One person	Away	Front	Partially concealed
13	Male 1, Female 1	Two people	Towards	Front	Exposed
14	Male 1, Female 1	Two people	Towards	Back	Exposed
15	Male 3, Female 2	Two people	Away	Front	Exposed
16	Male 3, Female 2	Two people	Away	Back	Exposed

* The 16 images are presented in Appendix C, Figure C.1.

Based on the results of Experiment 1, Experiment 2 was designed to further investigate only the face and hands, which one is more important on perceived safety evaluations (see Section 5.1). This involved varying the degree to which the face and/or the hands were concealed. Test participants were instructed to evaluate their perceived safety when observing test images of an approaching pedestrian. It was hypothesised that greater differences in evaluations caused by changes in concealment would suggest this was a more important visual cue. The predicted outcomes are presented in Table 3.7.

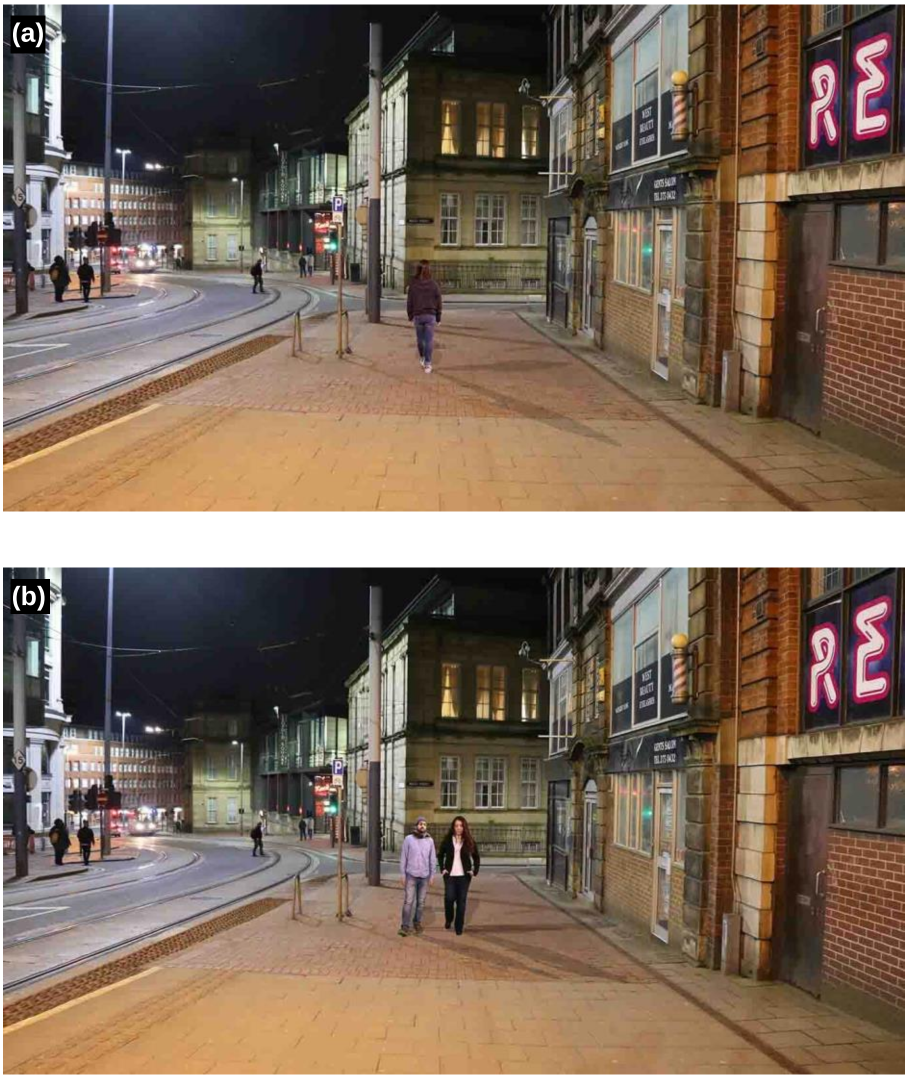


Figure 3.10. Examples from the test images used in Experiment 1. *(a)* Image (no. 11) shows Female 2 walking away in a front-lit condition with exposed face and hands. *(b)* Image (no. 13) shows two people (Male 1, Female 1) walking towards the camera in a front-lit condition with exposed face and hands.

Table 3.7. The tested cues and their predicted order of perceived safety.

Visual cues (of the approaching person)	**Predicted order of perceived safety evaluation of the four poses**			
	1 (safe)	**2**	**3**	**4 (less safe)**
Face concealment	Whole face is exposed	Face partially concealed – either top or bottom covered up		Whole face is concealed
Hand concealment	Hands exposed, either at the sides of the body or in front of the body		Hands concealed, either in the pockets or behind the body	
Mixed concealment*	Exposed face and hands	Exposed face while hands concealed or concealed face while hands exposed		Concealed face and hands

* Mixed concealment = face and hand concealment combined.

In Experiment 2, there were three test blocks of trials evaluating the influence of face concealment, hand concealment, and face and hand concealment combined (mixed concealment). In each block, the four actors portrayed the four levels of concealment presented in Table 3.8. For example, in face exposure variations, the actors were asked to conceal their face (part and whole) in different poses while their hands at their sides so that they were exposed; in hand exposure variations, the actors were asked to put their hands in different poses (exposure and concealed) while their faces remained fully exposed in all the images; in face and hand exposure variations, the same poses of face and/or hands were used in different combinations. The decision of obscuring the face and hands in Experiment 2 was made after sought with different methods such as pixelating and black colour (see Appendix B, Table B.1). The final method of concealment was chosen according to whether it was realistic; for instance, pixelation is not something people see on other pedestrians' faces.

Table 3.8. Example of the poses for each test block used in the experiment.

Block	Pose 1	Pose 2	Pose 3	Pose 4
Face concealment	 Exposed, i.e., not obscured by clothing	 Top part concealed, i.e., covered by wearing a hood and sunglasses	 Lower part concealed, i.e., covered by wearing a scarf	 Concealed, i.e., top and lower part of face covered by wearing hood, glasses, and scarf
Hand concealment	 Exposed, i.e., at the sides of the body	 Exposed, i.e., in front of the body	 Concealed, i.e., in the pocket	 Concealed, i.e., behind the body
Face and hand concealment combined (mixed concealment)	 Face and hands exposed, i.e., face not obscured by clothing, hands at the sides	 Face concealed and hands exposed, i.e., face is covered with hood, scarf, and sunglasses, hands at the sides of the body	 Face exposed and hands concealed, i.e., face not obscured by clothing, hands behind the body	 Face and hands concealed, i.e., face covered with hood and sunglasses, hands behind the body

Each test block in Experiment 2 was involved a combination of various face/hand exposures. This gave a total of 16 test images (for each block) captured with different variations (Table 3.9). Example images are presented in Figure 3.11. The entire set of test images for each block can be found in Appendix C, Figures C.2 to C.4.

Table 3.9. Variations of the 16 test images for each block used in Experiment 2.

Image no.*	**Gender**	**Face block****	**Hands block*****	**Mixed block**	
		Face exposure	**Hands exposure**	**Face exposure**	**Hands exposure**
1	Male 1	Exposed	Exposed	Exposed	Exposed
2		Top part concealed	Exposed	Concealed	Exposed
3		Lower part concealed	Concealed	Exposed	Concealed
4		Concealed	Concealed	Concealed	Concealed
5	Male 2	Exposed	Exposed	Exposed	Exposed
6		Top part concealed	Exposed	Concealed	Exposed
7		Lower part concealed	Concealed	Exposed	Concealed
8		Concealed	Concealed	Concealed	Concealed
9	Female 1	Exposed	Exposed	Exposed	Exposed
10		Top part concealed	Exposed	Concealed	Exposed
11		Lower part concealed	Concealed	Exposed	Concealed
12		Concealed	Concealed	Concealed	Concealed
13	Female 2	Exposed	Exposed	Exposed	Exposed
14		Top part concealed	Exposed	Concealed	Exposed
15		Lower part concealed	Concealed	Exposed	Concealed
16		Concealed	Concealed	Concealed	Concealed

* The 16 images are presented in Appendix C, Figures C.2 to C.7.
** The hands of the actors were fully exposed (at the sides) in all images.
*** The face of the actors was fully exposed in all images.

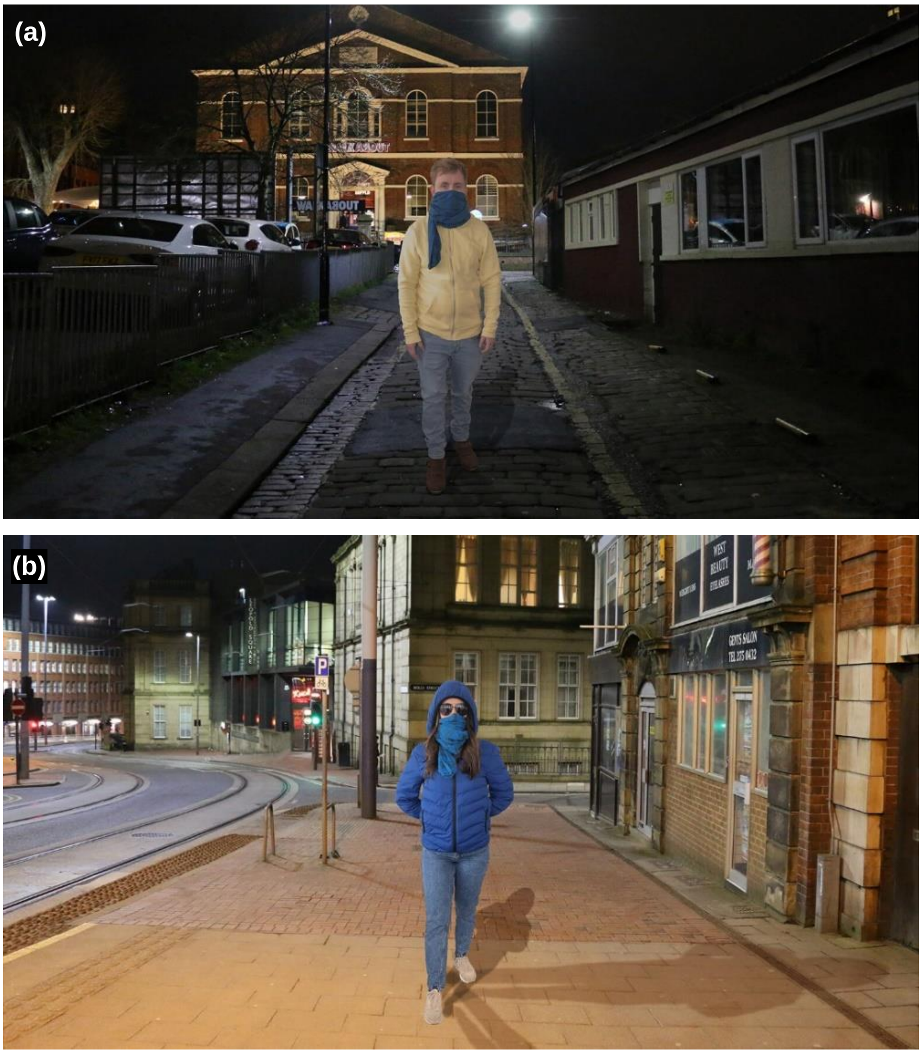


Figure 3.11. Examples from the test images used in Experiment 2. *(a)* Image from the face block (no. 6) shows Scene 2 with Male 2 displays Pose 3 (top part of face concealed). *(b)* Image from mixed block (no.12) shows Scene 1 with Female 1 displays Pose 4 (concealed face and hands).

3.4. General procedure

Two procedures were employed to evaluate the test images: Experiment 1 used category rating, while Experiment 2 used category rating and paired comparisons. The collected data provided a safety rating for each image, enabling variations in the photos to be compared. Using two procedures meant it was possible to assess the robustness of the results. If the conclusions drawn from evaluations of the same set of images using two different procedures tend to agree, then those conclusions may be considered more robust.

In the laboratory, the lights in the room were switched off and daylight from the window was obstructed using curtains, providing a dim ambient light level to simulate walking after-dark. In Experiment 1, the vertical illuminance at the observer's eye was 5 lux (mean across all test images), and the average horizontal illuminance on the desk in front of the participant was 3 lux (Figure 3.12). In Experiment 2, the vertical illuminance was 12 lux, and the average horizontal illuminance was 14 lux (Figure 3.13). At the start of each test session, 20 minutes was given to adapt to the low light level. During this time, a brief explanation of the test was given by the experimenter who then asked the participant to sign the consent form and read the participant information sheet (if needed). Examples of this information sheet and consent form can be found in Appendix E, Figures E.1 and E.2 (Experiment 1), and Appendix E, Figures E.3 and E.4 (Experiment 2).

Test participants in Experiment 1 were seated on a chair in front of the screen 0.50 m away from their eyes during the trials (Figure 3.12). The total time taken to conduct the experiment was approximately 35 minutes for each participant.

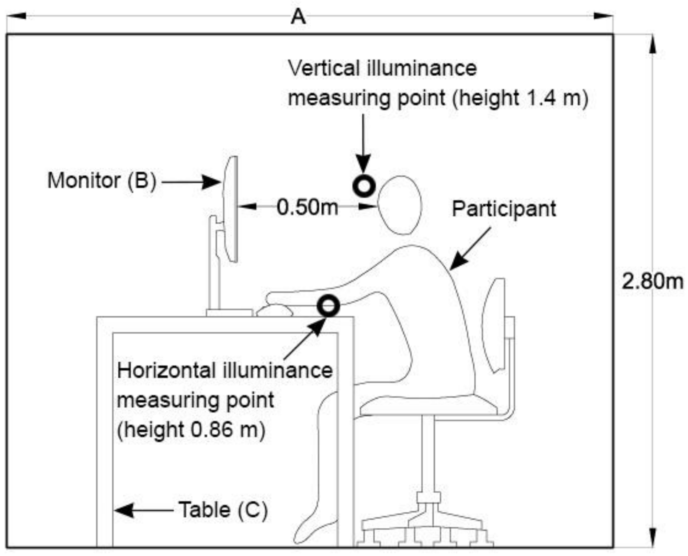


Figure 3.12. Diagram of the apparatus used in Experiment 1 which shows a section of a participant sitting on the chair while evaluating the images displayed on the monitor. Dimensions: A = laboratory length from front to back wall: 4.4 m; B = 27-inch monitor monitor dimensions: W = 0.57, H = 0.33; C = table dimensions: L = 1.4 m, W = 0.66 m, H = 0.86 m. The vertical centre of the presented image was aligned with the typical eye level of a sitting adult.

In Experiment 2, participants were required to walk upon a treadmill during trials to simulate walking after-dark (Figure 3.13). This is because our brain activity and the decisions we make differ when we walk as walking is associated with executive cognitive functioning, which increases the cognitive demand (Hausdorff et al., 2005; Yogev et al., 2008). The treadmill had a handrail nearby to allow recovery in the event of loss of balance. The treadmill speed was set to 2.5 km/h at the beginning of each trial, and test participants were instructed to change this, if required, to reach a comfortable walking speed. This is because walking uncomfortably (faster or slower) might cause a distraction and add difficulty to the task (Abernethy et al., 2002). Consequently, walking speeds during the experiment ranged from 1.5 to 3.5 km/h. When the participant was ready and comfortable on the treadmill, the experiment commenced, which included breaks of 2 minutes, if needed, or at any time was requested by the participant. It took approximately 50 to 60 minutes to complete all trials in each test session.

A recording of street sounds was played during trials, this is to enhance content validity and because the acoustic environment may affect mood and decisions (Fujikawa & Kobayashi, 2012). The recorded audio was mainly comprised the sound of wind with a few passing vehicles; it did not include any attention-attracting sounds such as animals, people shouting, phones ringing, or music. The recording lasted for 10 minutes and was looped to cover the whole test session.

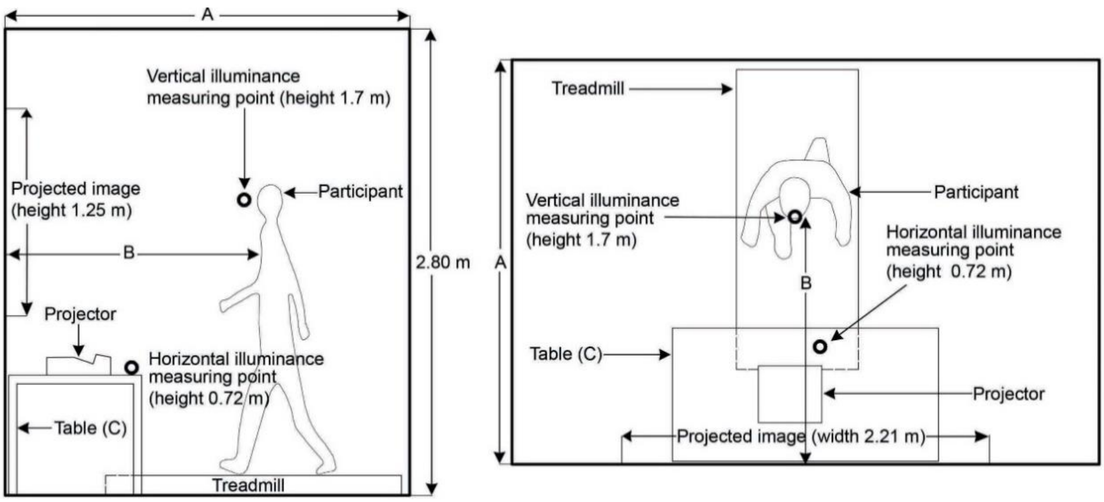


Figure 3.13. Diagram of the apparatus in Experiment 2: section *(Left)* and plan *(Right)*. Dimensions: A = 2.43 m (laboratory length from front to back wall). B = 1.5 m (from the projected image to the participant). C = table dimensions; L = 1.6 m, W = 0.80 m, H = 0.72 m. The vertical centre of the projected image was aligned with the typical eye level of a standing adult.

The experiments commenced with a demonstration and practice trials in each procedure to reinforce understanding of the instructions regarding use of the apparatus, and answering the same question in the main trials. The demonstration enabled the participant to test the experiment by clicking/pressing the buttons, confirming they were able to see the photograph, and ensuring they could understand the question presented. The practice tested whether the participants understood the task, and allowed the experimenter to answer any questions they may have had.

In Experiment 1, there were four practice trials that were the same size (0.60 x 0.34 m) as those used in the test trials, but displayed different actors and distances, as well as various background scenes. The actors' poses depicted a person approaching the camera with exposed faces and in a front-lit condition at three distances: 3 m, 7 m, and 9 m (Figure 3.14).

Figure 3.14. The four practice images used as practice trials in Experiment 1. These display different targets with different poses at various distances. *(a) and (b)* These show a female target and a male target, respectively, at a 3 m distance. *(c) and (d)* These show a male target at 7 m and 9 m, respectively.

In Experiment 2, the practice trials used four cartoon images in both procedures. These images represented a rating scale from extremely safe to extremely unsafe (Figure 3.15). The purpose of showing different levels of safety situations was to encourage the participant to think about safety ratings, similar to the main trials. The reason for choosing cartoon images was to validate the trials by exaggerating the scene, as the processing of emotions in cartoon faces is more accurate than real faces (Zhang et al., 2021); thus, there is no doubt when the image shows a safe or unsafe situation.

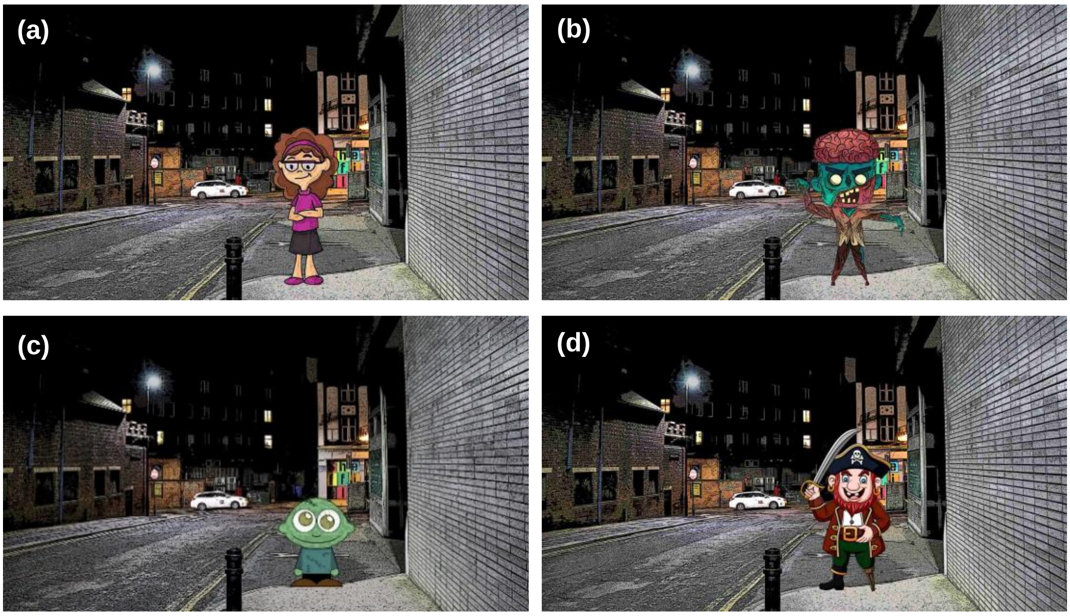


Figure 3.15. Images used in the practice trials in Experiment 2. *(a)* Image Considered to be a very safe situation. *(b)* Image considered to be an extremely unsafe situation. *(c) and (d) Images* show a middle situation either very/slightly safe or very/slightly unsafe, respectively.

The experiments presented the test images in a randomised order, using bespoke software based on a Python script that was written by Dr Chris Cheal (research associate in the Lighting Research Group at the University of Sheffield). This software automated the presentation process, including image randomisation, and recorded participants' responses onto an Excel spreadsheet. Then these responses were analysed using Microsoft Excel and IMB SPSS Statistic.

In Experiment 2, the order of the test blocks was fully randomised. The randomisation included the order of: the two procedures (category rating, paired comparisons), the three test blocks (face concealment, hand concealment, mixed concealment), the main 16 images displayed within a specific block, and the four practice images displayed. For these, methods of randomisation varied: image order randomisation was achieved using Python software; block order randomisation was attained by the participant rolling a dice at the start of each test session; and randomisation of the procedure order was performed manually by the experimenter using a checklist (Appendix D, Figure D.1). This checklist was created to check the randomisation order to ensure that each participant fulfilled all the requirements such as the adaptation time and vision tests, and to record additional data that might be needed such as the walking speed and the number of breaks.

3.4.1. Category rating for Experiments 1 and 2

At the beginning of a category rating block, the experimenter read the following statement aloud to the participants: *"Imagine you are walking alone at night along a road and see ahead another person, how safe would you feel?"* Next, a scenario was shown on the screen as text, and read aloud by the experimenter (Figure 3.16). Note: Experiment 2 used the same instruction, but with different answers:
"extremely unsafe │ very unsafe │slightly unsafe │slightly safe │ very safe │ extremely safe".

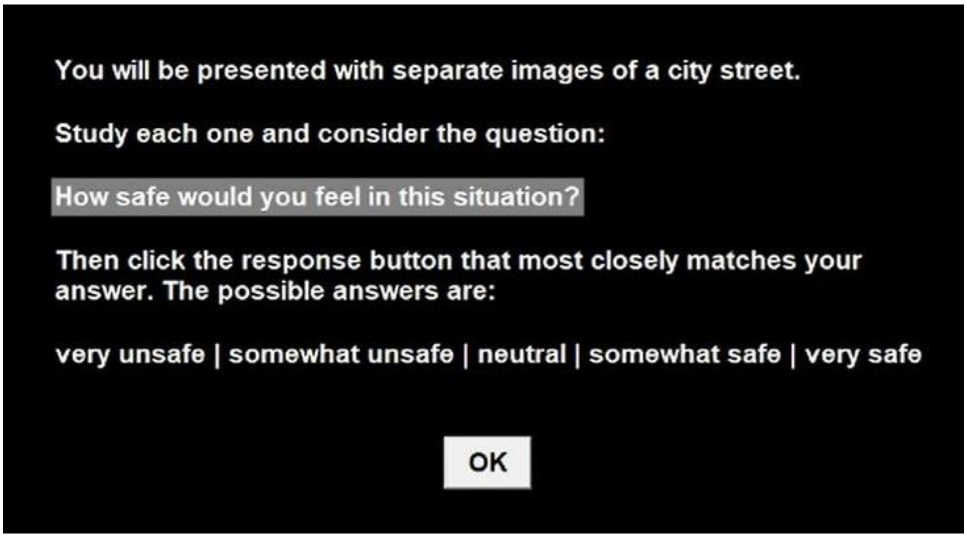


Figure 3.16. The instructions used in the category rating procedure in Experiment 1. This text was shown to the participant and read by the experimenter. Experiment 2 used the same instructions but with different answers.

Test participants were shown the 16 images one at a time. Each image was displayed for 0.5 seconds which is the typical duration of visual fixation upon other pedestrians (Fotios, Yang, & Uttley, 2015). After observing a given image, it was removed, leaving the question and a response scale for participants to complete. After providing a response, there was a two-seconds interval before the next image appeared for another evaluation. The interval was a blank black background in Experiment 1, and the same background scene with no target person in Experiment 2.

In Experiment 1, a 5-point response scale was used as suggested to meet the minimum required for the assumption of continuous data (Harpe, 2015). Participants responded to the given question on this scale by clicking (via the computer mouse) one of the appropriate choices shown on the monitor. A sample image is presented in Figure 3.17. Overall, there were 20 images to observe (16 test images + 4 practice), and approximately one minute was required to complete the task. Note: in both experiments, the category numbers (1 to 5) or (1 to 6) were not shown in the response scale, but subsequently numbered (1 to 5) or (1 to 6) for analysis.

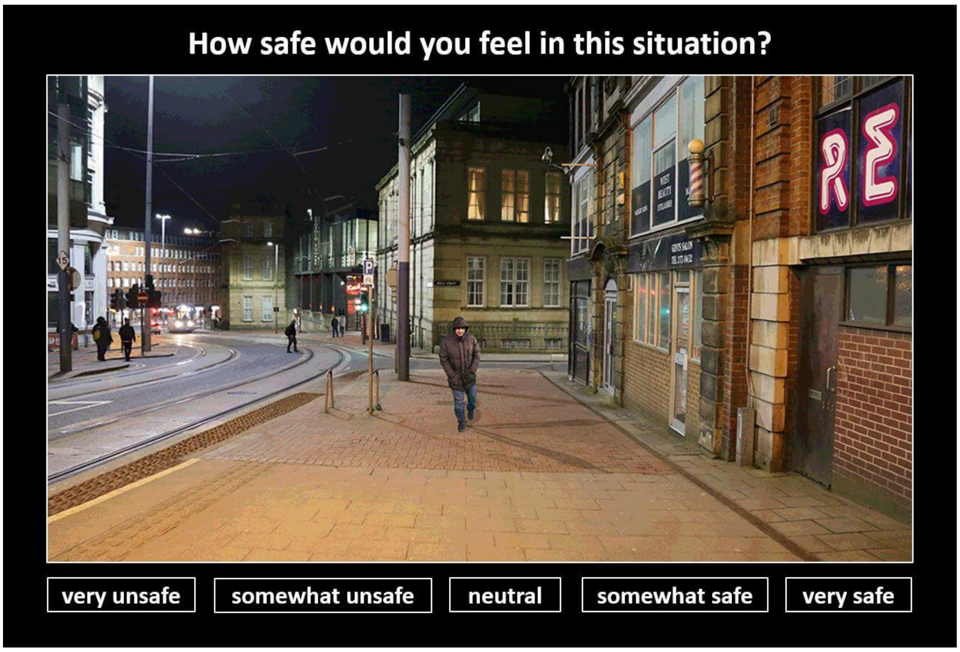


Figure 3.17. Example of a test image in the category rating procedure. This shows the question on the top and the 5-point responses at the bottom. The image in the middle is Image no. 2 which shows one person, Male 2, walking towards the observer, in a front-lit condition, with face and hands partially concealed; their face is partially covered by a hood and their hands are in their pockets.

In Experiment 2, a 6-point response scale was used because the even number of points avoids the middle option, and hence creates a forced choice (Fotios, 2019). Participants responded to the given question on this scale by pressing the appropriate button installed on the treadmill. A sample of a test image in the category rating is presented in Figure 3.18. In addition to the 16 test images per block, there were a further four repeated images chosen at random from the 16 test images within each block of trials, giving 20 images in total, observed in a randomised order. The repeated images were included to test the order effects on any unexpected differences (i.e. validation step) (Veitch et al., 2019); results of this are described in Section 5.2. Overall, 24 images were observed (16 test images + 4 practice + repeated images), and participants typically required three minutes to complete the evaluations of all the blocks.

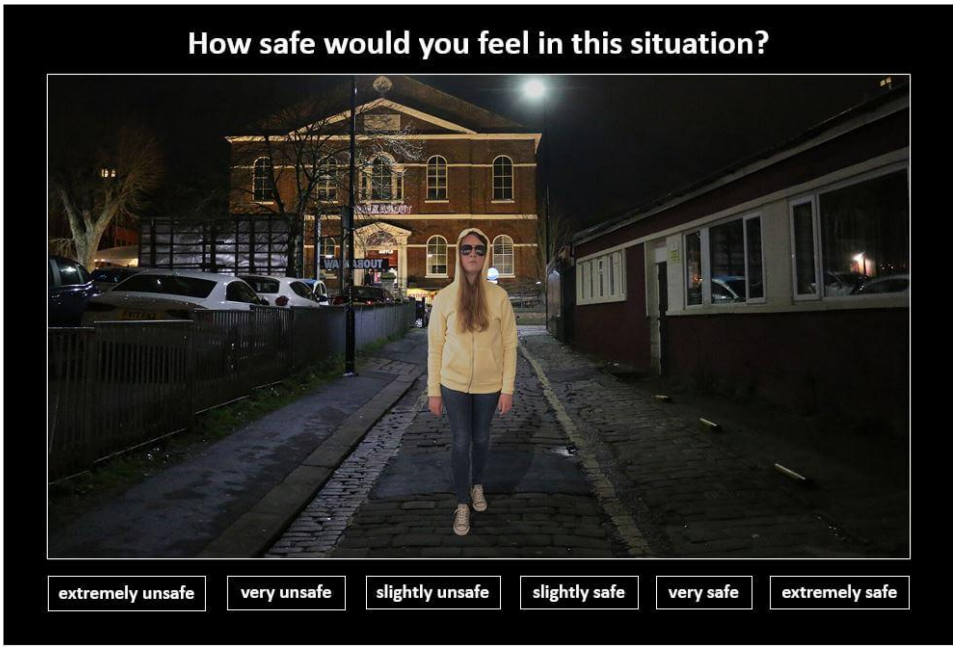


Figure 3.18. An example of a category rating test image shows the question on the top and the 6-point scale at the bottom. The image in the middle is Pose 2 in the face block, showing Female 2 with the top part of her face concealed with hood and sunglasses.

For an investigation of the effect of using different rating scales in safety evaluations. In Experiment 2, a different scale was used instead of the 5-point scale. This is because the middle choice (3 = neutral), which is seen as an easy choice for participants, might create a bias. Another reason for excluding the middle option (neutral) is that it does not provide a clear rating; is it safe or unsafe? It is likely to be a negative choice with a value of zero. Therefore, in Experiment 2, following the completion of the three blocks each with two procedures, a final task was conducted to explore the influence of different response scales. In this final task, the 16 test images of the face block were re-evaluated using the 5-point scale (the same as that used in Experiment 1) to be compared with the same images assessed on the 6-point scale, results of this are described in Section 5.2.3.

3.4.2. Paired comparisons for Experiment 2

At the beginning of a paired comparisons, a similar scenario to the category rating procedure was provided. Again, the experimenter read the following statement aloud to the participants: "*Imagine you are walking alone at night along a road and see ahead another person, how safe would you feel?"* Next, a scenario was shown on the screen as text, and read aloud by the experimenter (Figure 3.19).

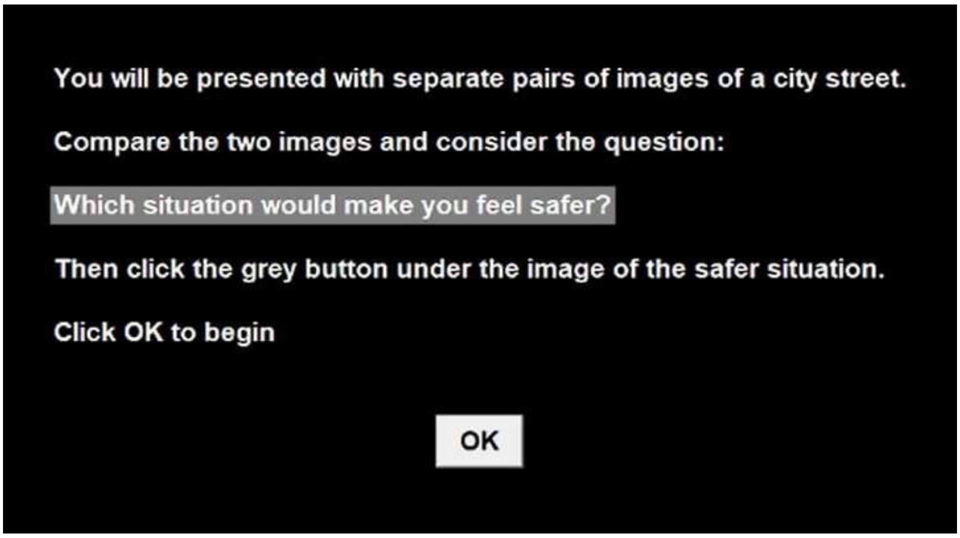


Figure 3.19. The instructions used in the paired comparisons procedure. This text was shown to the participant and read aloud by the experimenter.

Participants were instructed to choose what they perceived as the safer situation, either on the left or on the right image. This was a forced choice where the equally safe response (the option to choose 'same' or 'none') was not permitted. This was for two reasons: (1) it forces people to think and make an accurate decision that might reveal a difference; and (2) it is a criterion to start a method of analysis where responses are required to be 'yes' or 'no' (Dunn-Rankin et al., 2014).

All possible pairs of the 16 test images gives a total of 120 trials, these were shown to each participant, in a randomised order. Comparisons between specific pairs of images in a block demonstrated the effect of each pose. For example, comparing the face concealment in the face block between the Images 2 - 4, 6 - 8, 10 - 12, 14 - 16, where the faces were concealed or partly concealed against the Image 1, 5, 9, 13, where the faces were exposed revealed the effect of face covering; images can be seen in Appendix C, Figure C.2. The images were displayed side by side and viewed simultaneously for an unlimited duration. Following the onset of a given pair, there was a two-seconds delay before the response options appeared on screen and the response buttons became active. This was implemented as a precaution to avoid participants making a rapid response without first inspecting the images.

Responses were given by clicking one of the two buttons (left or right) installed on the treadmill. After giving a response, the target on both images disappeared, keeping only the background for a two-seconds interval before the appearance of the next target for evaluation. There were practice trials at the beginning, and null condition trials to check for position bias where the left and right scenes were identical. This meant that 135 pairs of images were observed in this procedure in each block (120 paired test images + 7 practice trials + 8 null conditions), and participants typically required 10 minutes per block to complete the evaluations. A sample from these test pairs of images is presented in Figure 3.20.

Figure 3.20. Example of a paired comparison test image. This is for the hands block, showing the question on the top, and the paired image in the middle. *Left:* image of Pose 2 shows Male 1 with hands exposed in front of the body. *Right:* image of Pose 4 shows Male 2 with hands concealed behind the body.

The paired comparisons procedure used one forced choice of pair images; therefore, to identify statistically significant differences between these pairs the scaling method of Dunn-Rankin Variance Stable Rank Sums (VSRS) was applied (Dunn-Rankin et al., 2014, p. 55). This significant difference is based on the difference between item sums and the critical range. Therefore, the critical range was calculated, which is based on the size of the sample (N) and the number of test items (K) required to produce an expected standard deviation and a value from the range distribution (Dunn-Rankin et al., 2014, p. 57), as formulated in the following equation; where K = the number of items and N = the sample size (number of judges).

$$E(S) = \sqrt{N(K)(K+1)/12}$$

Following this calculation, the critical range was found to be 37.1 for a sample of 44 participants. Next, the item sums were calculated, which were based on the total votes. The Dunn-Rankin analysis gave a total of 96 votes per participant rather than 120 (all possible comparisons of 16 images). This is because votes between images of different actors in the same pose were not counted. For instance, a paired image with different actors displaying the same pose (e.g., face concealed) was not counted because this pair did not reveal the effect of face. Based on this calculation, 96 votes was the maximum possible sum of range values per item of 132. The calculation of VSRS between items of a block is presented in Appendix K.

3.5. Participants

The experiments were ethically approved by the University of Sheffield (Experiment 1: reference: 018981 / 13 Jun 2018; Experiment 2: reference: 027364 / 03 July 2019). Following this, 32 participants for Experiment 1, and 44 for Experiment 2 were recruited with an equal balance of males and females. This sample was chosen according to Cohen's rule of thumb for sufficient effect size (for category rating data) (Cohen, 1992), and according to the minimum number of test participants necessary for the Dunn-Rankin method of statistical analysis (for paired comparisons data) (Dunn-Rankin et al., 2014, p. 59).

It was important to address questions of generalisability to the greater population, and to determine whether age and gender matter. Therefore, participants' age and gender were recorded to determine how well the collected data represent the whole population and society. The ages of the participants in Experiment 1 ranged between 18 to 54 years, with 25-34 years (69%) being the majority age group; in Experiment 2, the age range was between 18 to 34 years, with 25-34 years (88%) being the majority age group.

Participants produced written consent in accordance with ethical protocols. The research did not target people considered to be vulnerable. Each one received an incentive payment (£7 in Experiment 1, £10 in Experiment 2) for their contribution. This payment was made in recognition of the time and commitment provided by the participant, and to cover any travel or other expenses that may have been incurred in order to attend the experimental session.

Participants were given the consent form and the information sheet via email, which outlined all the details of the experiment before they came to the laboratory. If they agreed to participate, they received details of further requirements and how to access the laboratory. An example participant information sheet and consent form can be found in Appendix E, Figures E.1 and E.2 (Experiment 1) and Appendix E, Figures E.3 and E.4 (Experiment 2).

Each participant conducted the experiment on an individual basis. They were required to have reasonable eyesight, including normal colour vision and normal or corrected-to-normal visual acuity. This was screened at the beginning of the experiment using a Landolt ring acuity chart and the Ishihara colour test plates, illuminated by a D65 daylight-simulating lamp.

The Landolt C chart displays 12 rows, each of which contains five circles with a gap in four orientations (left, right, up, down). The size of the ring ranges from large to small, with the top first row displaying the largest rings and the bottom row displaying the smallest (Figure 3.21). Participants were required to indicate the correct direction of the gap at two meters either up, down, right, or left, while the experimenter marked the answers on a form. Normal visual acuity (minimum 6/6) was indicated if they were able to correctly identify gap direction in the top six rows (Kuo et al., 2011).

The Ishihara colour test plates were used to check for the presence of colour blindness. The booklet presents one or two digit numbers in a circle forming part of a series of circles in different colours and sizes (Figure 3.22). The participants passed the test if they were able to read the numbers on the plates. If they had trouble seeing red and green, they would not have been able to identify the numbers (Ishihara, 1987). No one was rejected on the basis of inadequate visual acuity or colour vision.

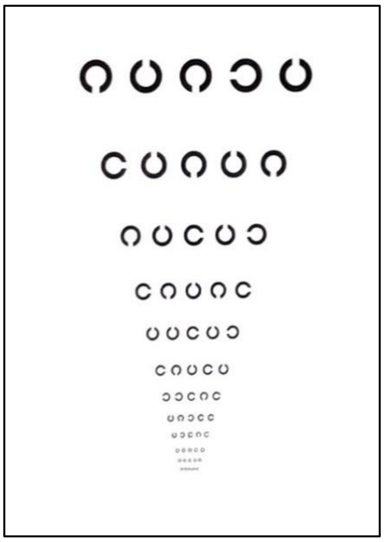


Figure 3.21. The Landolt ring acuity test sheet (printed in A4 size). The sheet was observed by participants from a distance of two meters.

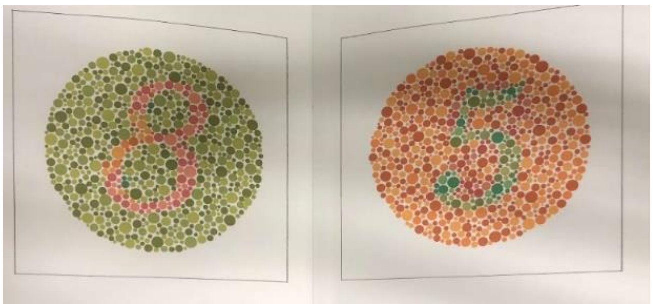


Figure 3.22. Example from the Ishihara colour test plates used in the experiments.

3.6. Limitations of Experiment 1

Implementation of Experiment 1 and subsequent analysis of the data suggested seven limitations that needed to be overcome in Experiment 2. First, the design of Experiment 1 was to use two males and two females to create the 16 test images, but a third male (Male 3) was accidentally used in four images (9, 10, 15, 16) which can be seen in Appendix C, Figure C.1. This extra male was only used in a walking away condition, and his face was never shown among the 16 test images. In Experiment 2, actors were chosen carefully with an equal number of males and females.

Second, errors were found when concealing the face and hands of some actors in the 16 test images. First, when concealing the face, two errors were identified: (1) Male 1 (in images 1, 5, 13, and 14) was wearing a winter hat while portraying an exposed face and hands. For a good pose, the head should not be concealed with any type of cover; (2) Male 2 and Female 2 (in images 2, 6, 4, and 8) were wearing a hood to conceal their faces, and they were trying to cover their face by moving the head down which meant the head direction varied between actors. For a good pose, the face should be completely hidden by a type of cover (e.g. wearing a sunglasses and/or face mask) and facing the camera, not other direction (e.g. the ground or left/right). Second, when concealing the hands, one error was identified: Female 1 (in images 4, 8, 13, and 14) was showing part of her hands while portraying a fully concealed hand pose. For a good pose, the hands should be completely hidden in pockets. Test images can be found in Appendix C, Figure C.1. It is unclear prediction how this might have affected the safety rating because face and hands were considered one variable and each test image was shown for 0.5 seconds, making it difficult for the participant to detect these differences when concealing the face and hands. However, to meet the variation criterion of the test variables and ensure consistency between them, actors in Experiment 2 were portrayed with the face facing towards the camera, while their hands were entirely hidden inside their pockets.

Third, the body size of some actors differed from the typical body size of British citizens. This was evident in four actors: the height of Female 1 (1.75 m) was above average (1.64 m); the weight of Male 2 (98 kg) was above average (79 kg); and the weight of two Chinese actors (Male 3/Female 2) (52 and 47 kg, respectively) was lower than average (79 and 70 kg, respectively). This information was presented in Table 3.2. It is not easy to predict how this might have impacted the safety rating, but there is a possibility that in the ‘walking away’ and ‘concealing the face and hands’ conditions, a tall female would be perceived as a male, or a short male as a female. Thus, in Experiment 2, actors were chosen whose body size was similar to the average body size of British citizens.

Fourth, for the background scene, the test participant was given a scenario of walking alone at night and seeing another person, and then asked how safe they would feel about his person. Therefore, the scene should support this by not showing any other pedestrians (except the actors) or distractions (e.g., vehicles or source of glare). Nevertheless, the scene included a few pedestrians far away at the back of the road, and an approaching tram with slight glare light (Figure 3.1). This could have affected the results of a safety rating in that a participant might have looked at the other persons rather than the test target; however, those extra pedestrians were very small and far away from the test actor which made them extremely difficult to notice, especially with an observation duration of 0.5 seconds. In Experiment 2, those extra pedestrians and the tram were removed from the scene (Figure 3.2).

Fifth, testing the exposure of the face and hands as one variable was not appropriate because, if a significant difference was suggested in the exposure of face and hands, it would not be clear which is more important. In Experiment 2, variations of face and hands were tested separately to enable a comparison, and to determine which has more of an effect.

Sixth, in test images, each of the five cues had two levels of variation (Table 3.6). To test all possible combinations would require 45 different images, but only 16 images were used in the experiment. This is because Experiment 1 deliberately concise by using a limited set of images rather than all possible combinations in order to explore the design of the method and assess whether the five tested cues (particularly the face) had an effect on perceived safety; based on this, Experiment 2 was able to use all possible combinations.

Seventh, the results of Experiment 1 indicated that the data were not normally distributed, which is recommended to use median values (see Section 4.2). However, sometimes reporting the median hides the differences between items, therefore, both the mean and the median were reported (see Section 4.3).

3.7. Summary

This chapter described the method of Experiments 1 and 2 conducted to investigate interpersonal evaluations of other pedestrians after-dark by comparing the relative importance of different visual cues identified in the literature (see Section 2.4). A combination of visual cues was established by generating a series of test images that show different actors embedded in background scene(s) after-dark. The actors portrayed the potential cues (gender, number of people, walking direction, light direction, exposure of face and hands) with variations between them. Test participants were required to evaluate the perceived feeling of safety for these images which were presented on a 27-inch PC monitor or projected onto a laboratory wall. The next chapter presents the results of Experiment 1 which aimed to explore interpersonal evaluation of pedestrians and investigated different visual cues.

Chapter 4. Safety Evaluation – Experiment 1: Results

4.1. Introduction

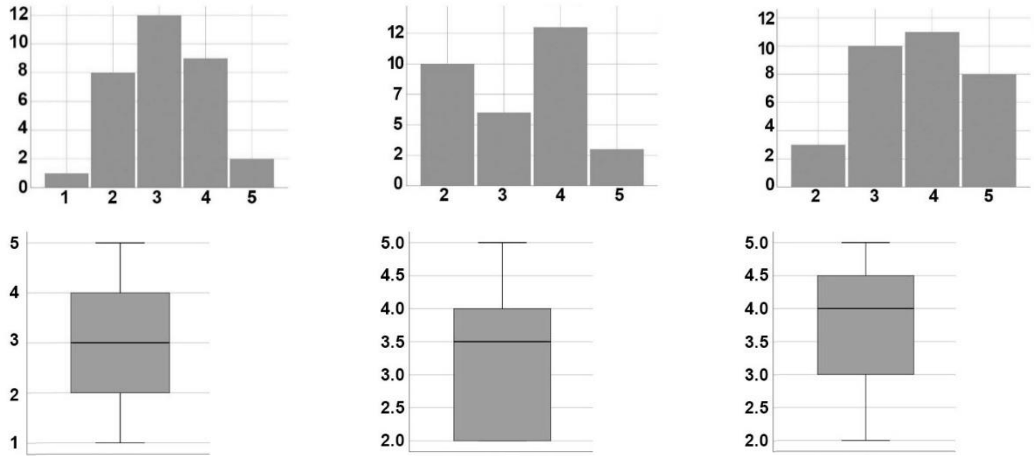
Chapter 3 described the methods employed in Experiments 1 and 2 to investigate the visual cues used for the interpersonal evaluation of other people. This chapter presents the results of Experiment 1 in which test participants evaluated the degree of safety associated with different pedestrians in test images manipulated to compare the influence of five visual cues: gender, number of people, walking direction, light direction, and the exposure of face and hands.

4.2. Distribution normality

The experimental results consisted of the ratings given by the 32 test participants to each of the 16 test images on a 5-point response scale where 1 = very unsafe and 5 = very safe. The raw results of Experiment 1 are reported in Appendix G, Table G.1.

Before summarising these data and conducting statistical analyses, it was important to determine whether the data were drawn from a normally distributed population or not. This comprised four steps: comparing the mean and median measures of central tendency, graphical representations of the data, measures of dispersion, and statistical tests (Appendix H, Table H.1). First, in comparing the mean and median, if the median value was within the range of the mean confidence level (CI 95%), then the decision was *yes* (normal); if not, then it was *no* (not normal). Second, graphical representations involved visual inspections of histograms and boxplot (Figure 4.1). Three decisions were possible: *yes*, *no*, and *near*. The final decision was *yes* if both were *yes*, *no* if both were *no,* or *near* if one was *yes* and one was *no*. If the latter, a decision was made according to how close the shape in the graphs was to normality. Third, measures of dispersion involved assessing skewness and kurtosis by calculating z-scores, where the values should be nearing a value of zero in a normal distribution (skewness: within ±0.5, kurtosis: within ±1.0) (Field, 2009, p. 138) . The final decision was *yes* if both tests were *yes*, otherwise it was *no*. Fourth, two statistical tests were performed: Kolmogorov-Smirnov and Shapiro-Wilks. The significance level in these tests was 0.05; therefore, if $p < 0.05$, the distribution was not normal. The final decision was *yes* if both were *yes*, otherwise it was *no*. Finally, for the overall assessment of normality, the variables (images) presented an acceptable normal distribution if there were at least three *yes* decisions within these four steps (Table 4.1).

This analysis did not suggest that the data were drawn from a normally distributed population. Therefore, the central tendency and distribution of each data sample were described using the median and inter-quartile range (IQR), and statistical analyses were conducted using non-parametric methods such as the Friedman test.



(a) (b) (c)

Figure 4.1. Example for testing data normality using the outcomes of graphical representations of the histograms (top) and boxplots (bottom). *(a)* Decision of *yes:* normal distribution, this is the outcome for Image 6. *(b)* Decision of *no*: not normal, this is the outcome for Image 2. *(c)* Decision of *near*, this is the outcome for Image 1.

Table 4.1. Example of final assessment of the normal distribution of the data; this displays an example of a decision of *no* (not normal) for Image 5, and *yes* (normal) for Image 6.

		Image 5	**Image 6**
Central Tendency	Mean 95% CI of mean* Median	3.31 2.89-3.74 3.00	3.09 2.75-3.44 3.00
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes
Graphical	Histogram Box Plot	Near Near	Yes Yes
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	Yes
Measures of dispersion	Skewness (within ±0.5) Kurtosis (within ±1.0)	-0.152 Yes -0.744 Yes	0.035 Yes -0.433 Yes
NORMALITY?	(Yes if both are yes)	Yes	Yes
Statistical tests Kolmogorov-Smirnov**	level of significance (not normal if p<0.05)	0.023 No	0.003 No
Shapiro-Wilks***	level of significance (not normal if p<0.05)	0.015 No	0.010 No
NORMALITY?	(Yes if both are yes)	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**Yes**

4.3. Results

Table 4.2 presents the median and the IQR of the safety ratings, in addition to the mean and the standard deviations for each test image (despite the data not being normally distributed). This is because the median values were extremely close to each other (3, 3.5 and 4) which may have hidden apparent differences. For example, 9 of the 16 test images (56%) had a median rating of 4, and the remaining 7 images (44%) had a median rating of 3 (only Image 2 rated 3.5). Consequently, the median ratings for both male and female targets was 4, which conceals the difference. In comparison, the difference between the mean ratings for the female targets was found to be 0.3.

As indicated in the results presented in Table 4.2, Image 6 was rated the least safe situation, and Image 3 the most safe; Image 6 showed one male walking towards the observer, with partially concealed face and hands in a backlit condition. By contrast, Image 3 showed one female walking towards the observer, with exposed face and hands in a front-lit condition (Figure 4.2).

Table 4.2. Safety ratings for each test image in Experiment 1; the 16 test images are ordered from least safe to most safe according to the mean ratings.

Image no.*	**Gender**	**Number of people**	**Walking direction**	**Light direction**	**Exposure of face and hands**	**Safety ratings**			
						Median	**IQR ****	**Mean**	**SD *****
6	M2	One	Towards	Back	Partially concealed	3.0	2	3.09	0.95
14	M1, F2	Two	Towards	Back	Exposed	3.0	2	3.19	0.88
12	F1	One	Away	Front	Concealed	3.0	2	3.22	1.02
10	M3	One	Away	Front	Concealed	3.0	2	3.25	0.97
2	M2	One	Towards	Front	Partially concealed	3.5	2	3.28	1.01
5	M1	One	Towards	Back	Exposed	3.0	2	3.31	1.16
8	F2	One	Towards	Back	Partially concealed	3.0	2	3.34	0.99
13	M1, F2	Two	Towards	Front	Exposed	4.0	1	3.72	0.94
1	M1	One	Towards	Front	Exposed	4.0	2	3.75	0.94
7	F1	One	Towards	Back	Exposed	4.0	1	3.88	0.78
4	F1	One	Towards	Front	Partially concealed	4.0	1	3.91	0.80
16	M3, F2	Two	Away	Back	Exposed	4.0	2	3.94	0.75
9	M3	One	Away	Front	Exposed	4.0	1	4.09	0.68
11	F2	One	Away	Front	Exposed	4.0	1	4.25	0.66
15	M3, F2	Two	Away	Front	Exposed	4.0	1	4.28	0.80
3	F1	One	Towards	Front	Exposed	4.0	1	4.31	0.73

* The 16 images are presented in Appendix C, Figure C.1.
** IQR = Inter-quartile range
*** SD = Standard deviation

Figure 4.2. The highest and lowest rated images. *(a)* The least safe situation rated for Image 6 shows one male walking towards the observer, with partially concealed face and hands in a backlit condition. *(b)* The most safe situation rated for Image 3 shows one female walking towards the observer, with exposed face and hands in a front-lit condition.

To check the proposed hierarchy of safety factors for pedestrians (see Section 3.3), the median (in addition to the mean) ratings across the 16 images was calculated, and each visual cue was analysed separately. Table 4.3 presents the results of the predicted safer situation for each visual cue and the difference between them by comparing the median ratings (and also the mean).

Table 4.3. Results of the predicted safer situation for each visual cue across the test images.

Visual cues	**Predicted less safe**	**Safety rating across indicated images**		**Difference between**		**Trend supports prediction**
				Median	**Mean**	
Gender	Female	Male: median = 4, mean = 3.5	Female: median = 4, mean = 3.8	0	0.3	Yes
Number of people	One person	One person: median = 4, mean = 3.6	Two people: median = 4, mean = 3.8	0	0.2	No
Walking direction	Away	Towards: median = 4, mean = 3.6	Away: median=4, mean=3.8	0	0.2	Yes
Light direction	Front	Back: median = 3.5, mean = 3.5	Front: median = 4, mean = 3.8	0.5	0.3	Yes
Exposure of face and hands	Exposed	Partially concealed: median = 3, mean = 3.3	Exposed: median = 4, mean = 3.9	1	0.6	Yes

The Friedman test revealed significant differences in safety ratings (p < 0.001) across all 16 images. Differences between pairs for each cue were analysed separately using the Wilcoxon test where p < 0.05 was significant (Field, 2009, p. 537).

4.3.1. Effect of gender

It was predicted that male target would lead to lower safety ratings (less safe) than female target (see Section 3.3, Table 3.5). The overall mean when the target was male was 3.5, and 3.8 when the target was female, which supports the prediction (Table 4.3). This suggests pedestrians feel less safe when a single male is approaching than when a single female is approaching.

To identify significant differences, the Wilcoxon test compared six pairs of images chosen so that factors other than gender were consistent (Table 4.4). These pairs were examined by comparing evaluations of images containing male target with images containing female target. In all pairs, the target was walking towards the observer. In three out of six cases, a significant effect ($p \leq 0.05$) (pairs 1, 2, 3) was identified with the female target being considered safer (Pair 1). In all three cases, the target was walking towards the observer, so gender should have been identifiable. In the remaining three cases (pairs 4, 5, and 6), although the results indicate a tendency for females to be rated as safer, the effect of gender was not significant ($p > 0.05$). In these three cases, gender identification may have been hindered due to the face being partially hidden by the hood (Pair 4) and the targets facing away from the observer (pairs 5 and 6). These results suggest that if we can see the gender of the person, we tend to consider females as being safer than males.

Table 4.4. Significant differences between gender.

Pair no.	Pairs (image no.)		Median dif.*	Variables in the images				Wilcoxon test	Sig.
	Male	Female		Number **	Walk ***	Light ****	Exposure *****		
1	1	3	1	One	Towards	Front	Exposed	p=0.001	Yes
2	2	4	0	One	Towards	Front	Partially concealed	p=0.007	Yes
3	5	7	0	One	Towards	Back	Exposed	p=0.005	Yes
4	6	8	0	One	Towards	Back	Partially concealed	p=0.201	No
5	9	11	0	One	Towards	Front	Exposed	p=0.244	No
6	10	12	0	One	Towards	Front	Partially concealed	p=0.618	No

* A median difference >0 indicates female targets were rated as being safer than male targets.
** Number of people
*** Walking direction
**** Light direction
***** Exposure of face and hands

4.3.2. Effect of number of people

It was predicted that being approached by two people would lead to lower safety ratings (less safe) than being approached by only one person. The overall mean when the target was single was 3.6, and 3.8 when the target was two people (Table 4.3). Approaching one person is therefore perceived as less safe than approaching two people.

To identify significant differences, the Wilcoxon test compared three pairs of images chosen so that factors other than number of people were consistent (Table 4.5). These pairs were examined by comparing evaluations of images containing two people with images containing only one person.

For matching conditions (same walking direction, light direction, and hand/face exposure), there were two one-person images for each two-person image, and hence ratings of the one-person images were averaged for each test participant. This provided three pairwise comparisons (13 v 1&3; 14 v 5&7; 15 v 9&11). In each case, the ratings suggest single people were considered safer than pairs, but this difference was significant only for the first ($p < 0.01$) and second ($p < 0.001$) pairs, not the third ($p = 0.62$).

Table 4.5. Significant differences between the number of people.

Pair no.	Pairs (image no.)		Median dif.**	Variables in the images				Wilcoxon test	Sig.
	One person*	Two people		Gender	Walk***	Light ****	Exposure *****		
1	1, 3	13	0	M, F	Towards	Front	Exposed	p=0.006	Yes
2	5, 7	14	1	M, F	Towards	Back	Exposed	p=0.001	Yes
3	9, 11	15	0	M, F	Away	Front	Exposed	p=0.617	No

* One is male and the other is female, with the average used in this calculation.
** A median difference >0 indicates two people were rated as being safer than one person.
*** Walking direction
**** Light direction
***** Exposure of face and hands

4.3.3. Effect of walking direction

It was predicted that target/s walking towards the observer would lead to lower safety ratings (less safe) than target/s walking away from the observer. The overall mean when the target/s were walking towards the observer was 3.6, and 3.8 when the target/s were walking away from the observer (Table 4.3). This suggests pedestrians feel less safe from an approaching person/s walking towards them than from a person/s walking away from them.

To identify significant differences, the Wilcoxon test compared six pairs of images chosen so that factors other than walking direction were consistent (Table 4.6). These pairs were examined by comparing evaluations of images containing target/s walking towards the observer with images containing target/s walking away from the observer. In the first three pairs, the differences were not significant ($p > 0.617$). In the latter three pairs, the differences were significant ($p < 0.01$), suggesting a greater feeling of safety when people are walking away from an observer rather than towards them.

Table 4.6. Significant differences between walking directions.

Pair no.	**Pairs (image no.)**		**Median dif.***	**Variables in the images**				**Wilcoxon test**	**Sig.**
	Toward	**Away**		**Gender**	**Number ****	**Light *****	**Exposure ******		
1	1	9	0	M	One	Front	Exposed	p=0.640	No
2	2	10	0	M	One	Front	Partially concealed	p=0.906	No
3	3	11	0	F	One	Front	Exposed	p=0.617	No
4	4	12	-1	F	One	Front	Partially concealed	p=0.006	Yes
5	13	15	0	M, F	Two	Front	Exposed	p=0.002	Yes
6	14	16	1	M, F	Two	Back	Exposed	p=0.002	Yes

* A median difference >0 indicates walk-away targets were rated as safer than walk-toward targets.
** Number of people
*** Light direction
**** Exposure of face and hands

4.3.4. Effect of light direction

It was predicted that target/s in a backlit condition would lead to lower safety ratings (less safe) than target/s in a front-lit condition. The overall mean when the target/s was in a backlit condition was 3.5, and 3.8 when the target/s was in a front-lit condition (Table 4.3). This suggests pedestrians feel less safe when someone is in a back-lit condition than when in a front-lit condition.

To identify significant differences, the Wilcoxon test compared six pairs of images chosen so that factors other than light direction were consistent (Table 4.7). These pairs were examined by comparing evaluations of images containing target/s in a backlit condition with images containing target/s in a backlit condition. In all six pairs, the front-lit target/s were considered safer than back-lit target/s, and this difference was significant in five cases ($p < 0.05$, pairs 1 and 6; $p < 0.01$, pairs 3, 4 and 5).

Table 4.7. Significant differences between light directions.

Pair no.	**Pairs (image no.)**		**Median dif.***	**Variables in the images**				**Wilcoxon test**	**Sig.**
	Front-lit	**Back-lit**		**Gender**	**Number ****	**Walk*****	**Exposure ******		
1	1	5	0	M	One	Towards	Exposed	p=0.016	Yes
2	2	6	0	M	Two	Towards	Partially concealed	p=0.206	No
3	3	7	0	F	One	Towards	Exposed	p=0.002	Yes
4	4	8	0	F	One	Towards	Partially concealed	p=0.002	Yes
5	13	14	-1	M, F	Two	Towards	Exposed	p=0.001	Yes
6	15	16	0	M, F	Two	Away	Exposed	p=0.044	Yes

* A median difference >0 indicates target/s in a backlit condition were rated as being safer than target/s in a front-lit condition.
** Number of people
*** Walking direction
**** Exposure of face and hands

4.3.5. Effect of exposure of face and hands

It was predicted that targets with face and hands partially concealed would lead to lower safety ratings (less safe) than targets with face and hands exposed. The overall mean when the target's face and hands were partially concealed was 3.3, and 3.9 when the target's face and hands were exposed (Table 4.3). This suggests pedestrians feel less safe when face and hands are concealed or partially concealed than when face and hands are exposed. This supported the predictions which were confirmed by the statistical analysis reported below.

To identify the significant differences, the Wilcoxon test compared six pairs of images chosen so that factors other than exposure of face and hands were consistent (Table 4.8). These pairs were examined by comparing evaluations of images containing target with exposed face and hands with images containing target with partially concealed face and hands. In all six pairs, the target with exposed face and hands were rated as being safer than those with partially concealed face and hands, and this difference was significant in five cases (p < 0.05, pair 2; p < 0.01, pairs 1, 4, 5, 6).

Table 4.8. Significant differences between the exposure of face and hands.

Pair no.	**Pairs (image no.)**		**Median dif.***	**Variables in the images**				**Wilcoxon test**	**Sig.**
	Exposed	**Partially concealed**		**Number ****	**Gender**	**Walk*****	**Light ******		
1	1	2	0	One	M	Towards	Front	p=0.007	Yes
2	3	4	0	One	F	Towards	Front	p=0.022	Yes
3	5	6	0	One	M	Towards	Back	p=0.254	No
4	7	8	0	One	F	Towards	Back	p=0.008	Yes
5	9	10	0	One	M	Away	Front	p=0.001	Yes
6	11	12	-1	One	F	Away	Front	p<0.001	Yes

* A median difference >0 indicates that target with partially concealed face and hands were rated as being safer than target with exposed face and hands.
** Number of people
*** Walking direction
**** Light direction

4.4. Summary

The aim of Experiment 1 was to explore and examine a range of visual potential cues used in the interpersonal evaluation of other pedestrians; specifically, how we decide whether or not it is safe to approach other people when walking after-dark. A total of 16 test images were observed on a PC screen and evaluated using a category rating procedure on a 5-point scale. The 16 images presented the following visual cues: gender, number of people, walking direction, light direction, and the exposure of face and hands.

All cues appeared to have the expected different effect on safety evaluations. It was confirmed that single males, face and hands partially concealed, a backlit condition, and walking towards the observer were all associated with a lower sense of safety. For the five tested cues, the mean difference and the statistical test suggested that regarding gender, females were considered safer than males; regarding the number of people, two people were considered safer than one person; regarding walking direction, walking away was viewed as safer than walking towards (but this was not supported by the statistical test); regarding light direction, the front-lit condition was considered safer than the backlit condition; regarding the exposure of face and hands, exposed face and hands were considered safer than partially concealed face and hands.

In conclusion, the factors that exerted the strongest influence on perceived safety was the exposure of face and hands because the difference between the two levels (exposed vs partially concealed) was greater than the differences for any of the other cues. However, the results were not sufficient to establish a reliable rank order for these factors due to the small differences between median/mean values. Therefore, there is no clear hierarchy for pedestrians' safety was found in this experiment.

Nevertheless, there are several limitations in this experiment which need to be addressed. First, some of the actors in the test images did not present an accurate covering of face and hands, and the size of the body (height and weight) of some actors differed substantially from the normal size of British citizens. Second, in the analysis, face and hands were tested as one variable, and average responses were reported using both the median and mean safety ratings. These limitations were discussed in detail in Section 3.6.

Chapter 5. Safety Evaluation – Experiment 2: Results

5.1. Introduction

The results of Experiment 1, presented in previous Chapter, suggest that exposure of the face and hands is an important cue when evaluating approaching pedestrians. Specifically, if face exposure is reduced by clothing (wearing a hood), and hand exposure is reduced by placing hands inside pockets, or if face and hand exposure are reduced by lighting (back-lit rather than front-lit), then the observer's degree of perceived safety is reduced. A limitation of Experiment 1 is that exposure of the face and hands was included as one variable – either both items were exposed or both were hidden – which meant it was not possible to determine whether one cue was more important than the other. Thus, Experiment 2 was conducted to further investigate the effect on safety evaluations of changes in face and/or hand exposure. This chapter presents the results and analyses of Experiment 2.

The test images in Experiment 2 were evaluated by 44 test participants using two procedures: category rating and paired comparisons. For the category rating procedure, a 6-point response scale was used where 1 = extremely unsafe and 6 = extremely safe. In the paired comparisons procedure, participants were asked to choose which of the two paired images presented the safer situation. Each procedure involved evaluating 16 test images (four actors each displaying four poses; see Table 3.8) in three blocks of trials: face, hands, and mixed. Two different background scenes (presented on Figure 3.2) were used in the test images with each scene evaluated by half of the sample (22 participants). The raw results for each test block and scene are reported in Appendix G, Tables G.2 to J.4.

5.2. Results: Category rating

The distributions of responses to each test image were checked before the main analysis to determine whether they were drawn from normally distributed populations. This was performed using the same statistical analyses described in Experiment 1 (see Section 4.2). These did not suggest the data were normally distributed for any of the three test blocks (Appendix H, Tables H.2 to H.4). Therefore, analyses were conducted using the median and IQR, and using non-parametric tests.

The results were first analysed to check for order bias. This was undertaken by comparing the responses for two identical images evaluated in repeated trials. Each of the 44 participants evaluated four repeated pairs in each of the three blocks, giving a total of 528 null condition trials across all blocks. In 352 (67%) of these trials, the same rating was given in both evaluations of the same test image. Of the remaining 176 trials, the first and second evaluations led to a safer rating on similar numbers of trials (86 safer on first trial, 90 safer on second trial).

Table 5.1 presents these null condition evaluations separate for each scene in a block, alongside the significant differences between first and second ratings. These differences were analysed using the Wilcoxon test where $p \leq 0.05$ indicates a significant difference (Field, 2009, p. 537). In all cases, the differences between first and second evaluations were not significant (p = 0.18 to p = 0.95), suggesting that any bias in the current results due to an order effect was negligible. In any situation, the test images were presented in a randomised order to offset order effects (Veitch et al., 2019). Subsequent analyses utilised ratings of the first observation of the repeated images. This ensured consistency with the other images, which used the first (and only) evaluation of each image.

Table 5.1. Results of the repeated trials analysis for each test block and background scene, showing the number of identical ratings (zero difference), second ratings as safer, and second ratings as less safe. The sample size in each scene was n = 22.

Test block	**Face**		**Hands**		**Mixed**	
Background scene*	**S1**	**S2**	**S1**	**S2**	**S1**	**S2**
No. of pairs with zero difference	50	65	61	56	60	60
No. of pairs safer on second trial	18	10	18	15	15	14
No. of pairs less safe on second trial	20	13	9	17	13	14
Significance of difference between first and second ratings of the same image (Wilcoxon test)	p=0.90	p=0.53	p=0.18	p=0.65	p=0.95	p=0.85

* S1 = Scene 1; S2 = Scene 2

Each of the four target actors displayed four poses in each block which were embedded into two different background scenes. For each image of block and scene, the median rating, the difference between the median ratings, and inter-quartile range (IQR) across all actors are presented in Table 5.2. Graphs of the median ratings for each actor and scene in a block can be found in Appendix I, Figures I.1 to I.3. To test for significant differences between the evaluations given in each scene, the Mann-Whitney test for independent samples was used because each scene was evaluated by a different sample of 22 people. For each combination of block and image, $p < 0.05$ indicates a significant effect. There were no significant differences between the evaluations given in each scene for all three blocks (Table 5.2). Therefore, the data were reorganised as one set of 44 test participants and the factor of different scenes was ignored.

Table 5.2. Median evaluation of perceived safety for each image in each combination of background scene and block; also presented the difference in median rating between Scenes 1 and 2, inter-quartile range (IQR), and significance of differences assessed using the Mann-Whitney test**

Image no.	**Face block**					**Hands block**					**Mixed block**				
	Median in scene		**Dif.***	**IQR**	**Mann-Whitney test****	**Median in Scene**		**Dif***	**IQR**	**Mann-Whitney test****	**Median in Scene**		**Dif.***	**IQR**	**Mann-Whitney test****
	S1	**S2**				**S1**	**S2**				**S1**	**S2**			
1	5.00	5.00	0.0	1	*p*=0.651	5.00	5.00	0.0	1	*p*=0.824	5.00	5.00	0.0	1	*p*=0.750
2	3.00	3.00	0.0	2	*p*=0.463	4.00	4.00	0.0	2	*p*=0.807	2.00	2.00	0.0	1	*p*=0.867
3	3.00	3.00	0.0	2	*p*=0.642	4.00	4.00	0.0	2	*p*=1.000	4.00	4.00	0.0	2	*p*=0.760
4	1.50	1.50	0.0	2	*p*=0.980	4.00	3.00	1.0	2	*p*=0.149	1.50	2.00	0.5	1	*p*=0.566
5	5.00	5.00	0.0	2	*p*=0.758	4.00	5.00	1.0	1	*p*=0.384	5.00	5.00	0.0	2	*p*=0.404
6	4.00	3.00	1.0	1	*p*=0.284	5.00	4.00	1.0	1	*p*=0.609	2.00	2.00	0.0	1	*p*=0.693
7	3.00	3.00	0.0	2	*p*=0.367	4.00	4.00	0.0	1	*p*=1.000	4.00	4.00	0.0	2	*p*=0.393
8	2.00	1.50	0.5	2	*p*=0.449	4.00	4.00	0.0	1	*p*=0.696	2.00	2.00	0.0	1	*p*=0.227
9	5.00	5.00	0.0	1	*p*=0.269	6.00	5.00	1.0	1	*p*=0.380	6.00	6.00	0.0	1	*p*=0.452
10	4.50	4.00	0.5	2	*p*=0.128	5.00	5.00	0.0	2	*p*=0.075	3.00	2.50	0.5	2	*p*=0.808
11	4.00	4.00	0.0	2	*p*=0.185	5.00	5.00	0.0	1	*p*=0.950	5.00	5.00	0.0	2	*p*=0.608
12	3.00	2.00	1.0	1	*p*=0.389	5.00	4.00	1.0	1	*p*=0.080	2.00	2.00	0.0	2	*p*=0.670
13	5.00	5.00	0.0	1	*p*=0.441	5.00	5.00	0.0	2	*p*=0.951	5.50	5.00	0.5	1	*p*=0.328
14	4.00	3.00	1.0	1	*p*=0.576	5.00	5.00	0.0	2	*p*=0.643	2.50	3.00	0.5	1	*p*=0.670
15	3.50	4.00	0.5	1	*p*=0.696	5.00	5.00	0.0	1	*p*=0.676	4.00	4.00	0.0	2	*p*=0.439
16	2.00	2.00	0.0	2	*p*=0.912	4.00	4.00	0.0	2	*p*=0.567	2.00	2.00	0.0	2	*p*=0.442

* Difference in median rating between Scenes 1 and 2
** $p < 0.05$ indicates a significant effect.

5.2.1. Analysis of test blocks

Within each block, there were 16 test images. Analysis of the evaluations for these images using the Friedman test revealed significant differences ($p < 0.001$) for all three blocks. For a given pose, the Friedman test identified significant ($p < 0.001$) differences between the four actors for all three blocks. For a given actor (and for all actors considered together), the Friedman test indicated significant ($p < 0.001$) differences between the four poses for all three blocks. For a given actor, differences between each pair of poses were analysed separately using the Wilcoxon test (Appendix J, Tables J1 to J3). These differences were not consistent between actors, and subsequent results will discuss these differences separately for each block.

A difference which is statistically significant does not mean it is of practical importance. Substantive significance is assessed according to the size of the effect (Field, 2009, p. 56). Thus, the effect size was estimated using Cohen's *d*, calculated between all combinations of the four poses.

For the face block, the median ratings for an image being considered safer are presented in Figure 5.1 which shows the distribution of ratings as safer for each pose of an actor. In general, a person was considered safer when the face was exposed, and females were considered safer than males. Poses with either the bottom or the top of the face concealed exhibited similar trends.

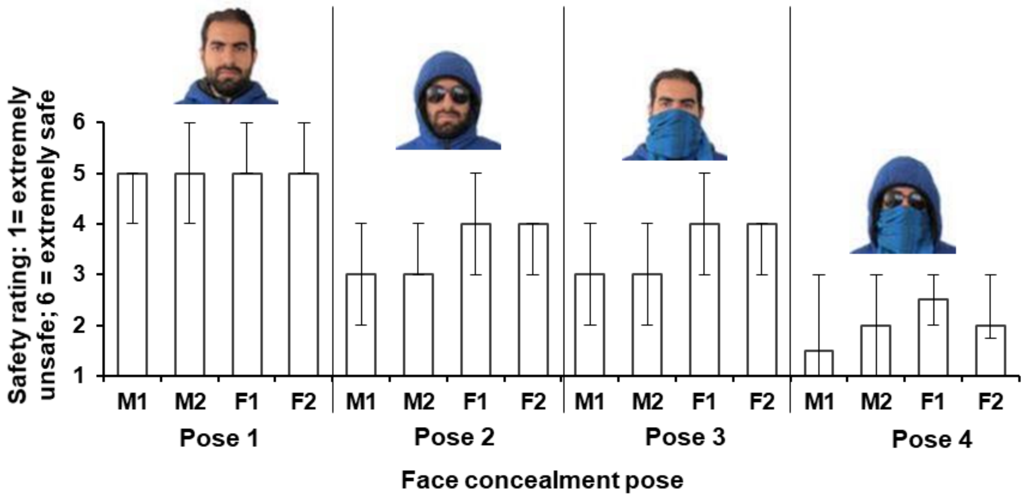


Figure 5.1. Median ratings of safety for each actor in a pose for the face block; Error bars denote the inter-quartile range. Information about each actor were illustrated in Table 3.3.

Figure 5.2 depicts the median of the evaluations for all actors in the same face pose. Increases in face concealment led to a gradual reduction in ratings of safety (from 5 to 2). Hence, the rating for Pose 1 (face exposed) was significantly higher (5) than for any other pose; the rating for Pose 4 (face concealed) was significantly lower (2) than for any other pose. The same rating (5) was found for Poses 2 and 3 (top part concealed, bottom part concealed).

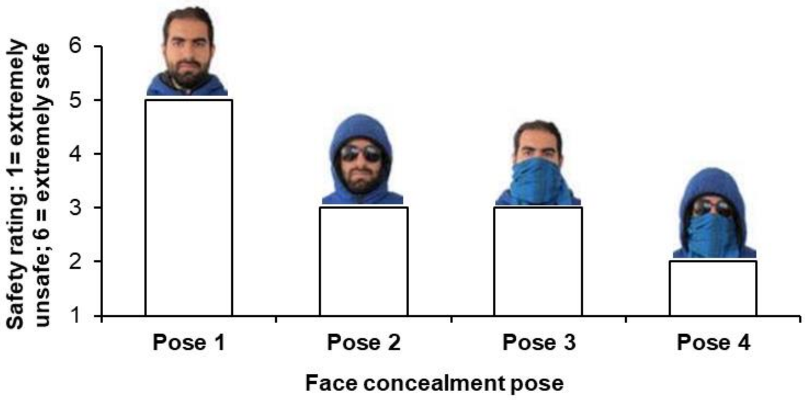


Figure 5.2. Median safety rating for each pose across the four actors in the face block.

Table 5.3 presents the significance of differences between the face poses and the corresponding effect size. There were no significant differences between evaluations of Poses 2 and 3. The differences indicated a large effect, except for between Poses 2 and 3 where the effect size was negligible.

Table 5.3. The significance of differences and the corresponding effect size between the face poses in the face block. Differences were tested using Friedman and were consistent for the four actors except where shown. The poses were illustrated in Figure 5.2.

Block		**Significance of difference**			**Effect size****		
		Pose 2	**Pose 3**	**Pose 4**	**Pose 2**	**Pose 3**	**Pose 4**
Face	**Pose 1**	p<0.001	p<0.001	p<0.001	Large	Large	Large
	Pose 2	---	n.s*	p<0.001	---	None	Large
	Pose 3	---	---	p<0.001	---	---	Large

* For actors M1, F1, F2, $p > 0.49$, for actor M2, $p = 0.091$.
** Cohen's *d*: r = 0.1 (small effect); r = 0.3 (medium effect), r = 0.50 (large effect).

For the hands block, the median ratings for an image being considered safer are presented in Figure 5.3, which shows the distribution of rating as safer for each pose by an actor. Generally, a person was considered less safe when their hands were located behind the body, and females were considered safer than males. The three poses 1, 2, and 3 (hands at side, in front, inside pocket) exhibited similar trends.

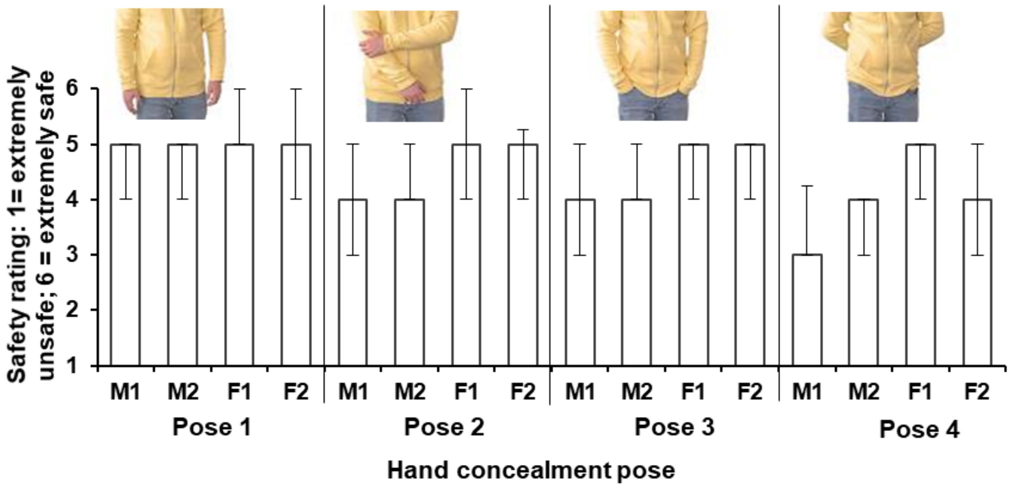


Figure 5.3. Median ratings of safety for each actor in a pose for the hands block; Error bars denote the inter-quartile range. Information about each actor were illustrated in Table 3.3.

For the hands block, Figure 5.4 presents the median of the 44 evaluations for all four actors using the same hand pose. Changes in hand concealment led to smaller changes in safety evaluations than changes in face concealment. The same median rating (5) was found for Poses 1, 2, and 3 (hands at side, in front, inside pocket); the lower median (4) was only for pose 4 (hands behind the back).

Table 5.4 presents the significance of differences between the hand poses and the corresponding effect size. Comparison of Poses 1 and 3 indicated a significant difference with a large effect; however, there were no significant differences between other comparisons (1 vs 2 and 2 vs 3) and the effect sizes were small and negligible. Pose 4 was significantly different between all poses with a large effect with Poses 1 and 2, and medium effect with Pose 3. Note: this medium effect presented the hands concealed in a similar way (in pockets and behind back), which suggests that hands behind the body (Pose 4) are perceived as less safe than other hand locations.

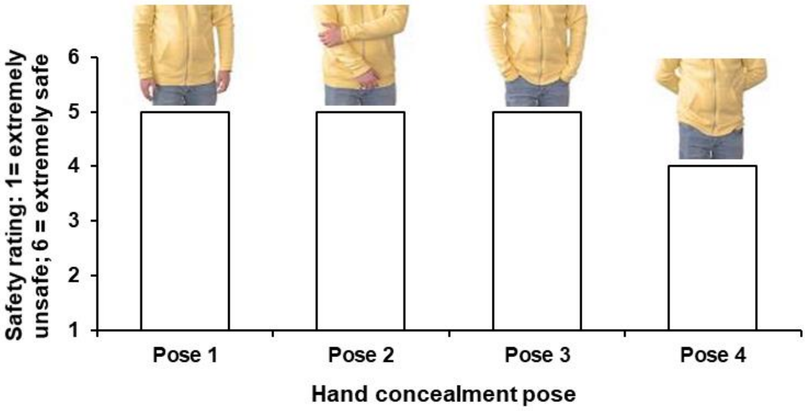


Figure 5.4. Median safety rating for each pose across all actors in the hands block.

Table 5.4. The significance of differences and the corresponding effect size between the hand poses in the hands block; differences were tested using Friedman and were consistent for the four actors except where shown. The poses were illustrated in Figure 5.4.

Block		**Significance of difference**			**Effect size*******		
		Pose 2	**Pose 3**	**Pose 4**	**Pose 2**	**Pose 3**	**Pose 4**
Hands	**Pose1**	n.s*	*p*<0.05**	*p*<0.001	Small	Medium	Large
	Pose2	---	n.s.	*p*<0.002	---	None	Large
	Pose3	---	---	*p*<0.05***	---	---	Medium

** M1, *p* = 0.046; M2, *p* = 0.35; F1, *p* = 0.044; F2, *p* = 0.243: overall conclusion, not significant.
*** M1 *p* = 0.027, F1, F2 *p* < 0.01, M2 *p* = 0.195.
**** M1, M2, *p* < 0.001; F2 *p* = 0.007, F1 *p* = 0.066.
***** Cohen's *d*: r =0.1 (small effect); r = 0.3 (medium effect), r = 0.50 (large effect).

For the mixed block, the median ratings for an image, being considered safer, are presented in Figure 5.5 which shows the distribution of rating as safer for each pose of an actor. In general, a person was considered less safe if the face and the hands were concealed, and females were considered safer than the males. Poses with either face concealed and hands exposed or, conversely, face exposed and hands concealed exhibited similar trends.

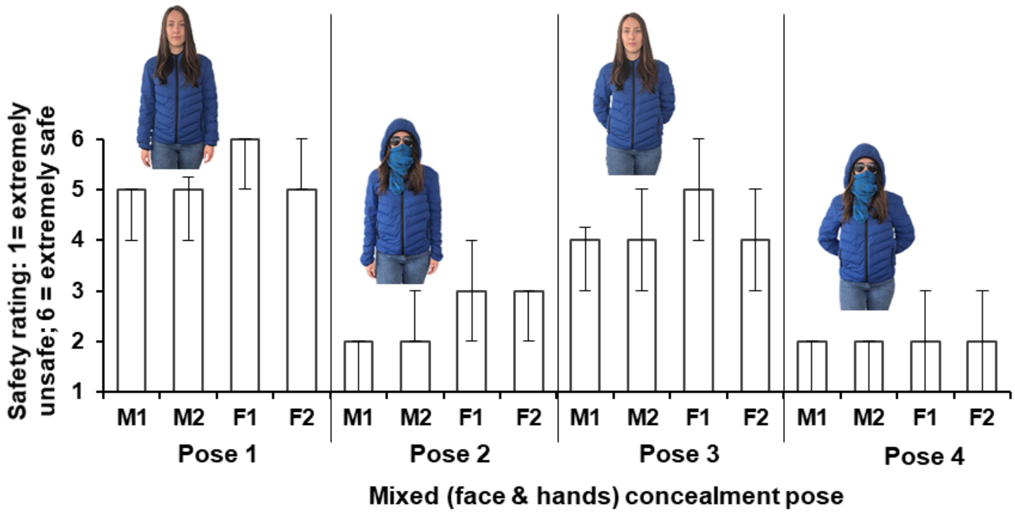


Figure 5.5. Median ratings of safety across for each actor in a pose for the mixed block; Error bars denote the inter-quartile range. Information about each actor were illustrated in Table 3.3.

For the mixed block, Figure 5.6 presents the median of the 44 evaluations for all four actors in the same mixed pose. The rating for Pose 1 (face and hands exposed) was significantly higher (5) than for other poses; a similar rating (4) was given for Pose 3 (face exposed and hands concealed). The rating for Poses 2 (face concealed and hands exposed) and 4 (face and hands concealed) were significantly lower (2). In other words, for the Poses 1 and 3 where the face was exposed (note: hands were concealed in pose 3), the safety rating was significantly higher (i.e., more safe) than the Poses 2 and 4 where the face was concealed (note: hands were concealed in Pose 4).

Table 5.5 presents the significance of differences between the mixed poses in the mixed block and the corresponding effect size. There were significant differences between evaluations of all poses with a large effect, except for Poses 2 and 4 where there was a medium effect. Thus, the effect of concealing the hands in Poses 1 and 3 has less impact on the safety rating than that of concealing the face in Poses 3 and 4. These results were similar to those were found in the face block and the hands block, where hand exposure was revealed to have less influence on safety than face exposure.

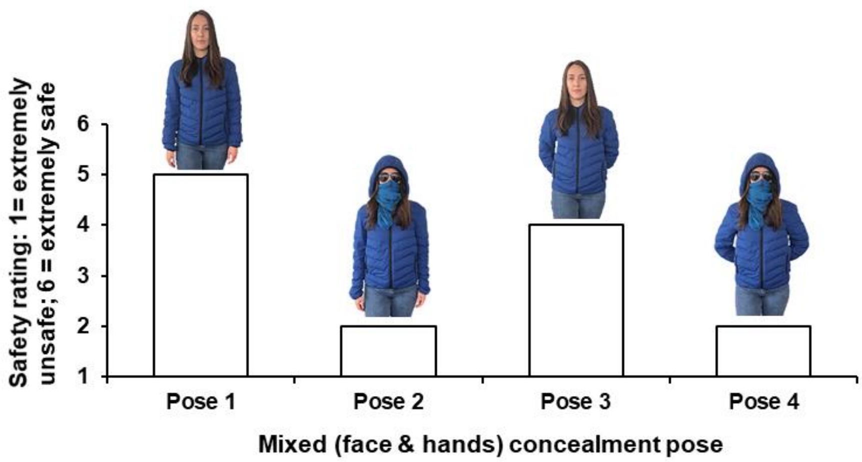


Figure 5.6. Median safety rating for each pose across all actors in the mixed block.

Table 5.5. The significance of differences and the corresponding effect size between the mixed poses in the mixed block; differences were tested using Friedman, and were consistent for the four actors except where shown. The poses were illustrated in Figure 5.6.

Block		**Significance of difference**			**Effect size****		
		Pose 2	**Pose 3**	**Pose 4**	**Pose 2**	**Pose 3**	**Pose 4**
Mixed	Pose 1	*p*<0.001	*p*<0.001	*p*<0.001	Large	Large	Large
	Pose 2	---	*p*<0.001	*p*<0.05*	---	Large	Medium
	Pose 3	---	---	*p*<0.001	---	---	Large

* M1 $p = 0.13$, M2 $p = 0.052$, F1, F2 $p < 0.01$.
** Cohen's *d*: r = 0.1 (small effect); r = 0.3 (medium effect), r = 0.50 (large effect).

5.2.2. Analysis of identical images rated in different blocks

Three identical images were evaluated in different blocks. Table 5.6 depicts the poses in these images, and in which blocks they were used. The results of these evaluations were compared in order to determine whether these poses were similarly evaluated.

Table 5.6. Example of the poses rated in more than one block.

Pose No.	**Pose 1**	**Pose 2**	**Pose 3**
Condition	 Face and hands exposed	 Face concealed, hands exposed	 Face exposed, hands concealed
Rated in blocks	Face, Hands, Mixed	Face, Mixed	Hands, Mixed

Table 5.7 presents an analysis using the median safety rating for each actor in a same pose for each test block. The significant differences between these evaluations were assessed using either Friedman or Wilcoxon. The choice of test depends on the number of variables under investigation, where Friedman requires three variables, and Wilcoxon requires two variables. For actor M1, there were no significant differences between the ratings in different blocks for all three poses. For actors M2, M3, and M4, there were significant differences between the ratings in different blocks ($p < 0.03$) in Pose 1 and Pose 3, but not for Pose 2 (according to Wilcoxon test).

Table 5.7. Significance of differences between the identical images rated in different blocks for each actor. The poses were illustrated in Table 5.6.

Actor reference*	Pose	Median rating for each actor in block:			Test of difference	*p*-value	Sig.
		Face	Hand	Mixed			
M1	1	5.00	5.00	5.00	Friedman	p=0.167	No
	2	1.50	---	2.00	Wilcoxon	p=0.794	No
	3	---	3.00	4.00	Wilcoxon	p=0.333	No
M2	1	5.00	5.00	5.00	Friedman	p>0.001	Yes
	2	2.00	---	2.00	Wilcoxon	p=0.617	No
	3	---	4.00	4.00	Wilcoxon	p=0.03	Yes
F1	1	5.00	5.00	6.00	Friedman	p>0.001	Yes
	2	2.50	---	3.00	Wilcoxon	p=0.557	No
	3	---	5.00	5.00	Wilcoxon	p=0.001	Yes
F2	1	5.00	5.00	5.00	Friedman	p=0.001	Yes
	2	2.00	---	3.00	Wilcoxon	p=0.176	No
	3	---	4.00	4.00	Wilcoxon	p>0.001	Yes

* Information about each actor can be found in Table 3.3.

Figure 5.7 presents the median rating across the four actors in the same pose rated in two or three test blocks. The same rating was found for these poses, where Pose 1 rated 5 (very safe), Pose 3 rated lower 4 (slightly safe), and Pose 2 rated lower 2 (very unsafe). The significance of the differences between these poses is detailed in Table 5.8. Poses 1 and 3 were significant but Pose 2 was not. These poses (images) were exactly the same but yielded slightly different results; this could be related to the order effect discussed in the next point.

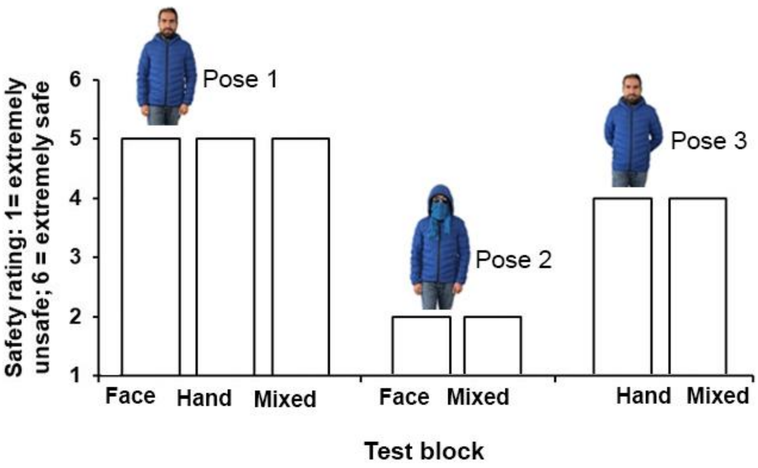


Figure 5.7. Safety rating across all the actors for each pose rated in different blocks.

Table 5.8. Significance of differences between the identical images in each block across all the actors. The poses were illustrated in Table 5.6.

Pose	**Median rating across all actors in block:**			**Test of difference**	***p*-value**	**Sig.**
	Face	**Hands**	**Mixed**			
1	5	5	5	Friedman	*p*=0.056	Yes*
2	2	---	2	Wilcoxon	*p*=0.208	No
3	---	4	4	Wilcoxon	*p*=0.017	Yes

* *p* = 0.05 is the normal threshold; because 0.056 is very close to the threshold, it is assumed to be significant.

In all blocks, the 16 images were presented in a random order. However, it is possible that there may have been an order effect in which the evaluation of a particular image was influenced by evaluations of preceding images (Fotios, 2019). Thus, the first image evaluated in a block would not be affected by order bias. Three analyses were conducted with different levels of order effect (applied for the face block only).

The first analysis used evaluations of the first images; this reduced the data sample from n = 44 (44 x 16) to n = 44 (44 x 1). Figure 5.8 presents the median ratings of perceived safety for each pose of face concealment across all actors; these results are identical to those of the complete data set (Figure 5.2). This suggests that rating an image at first observation does not change according to the order (e.g., third, fifth, or tenth). Nevertheless, these data were not completely free of order bias as the evaluations may have been influenced by evaluations of the hands and/or mixed targets in the rating scale procedure, as well as the evaluation from the paired comparison procedure when the face block was not the first to be employed. This means that all the images in this analysis were the first to be evaluated in the face block, but might be rated in one of the procedures of the hand and/or mixed block. The next analysis examined this effect.

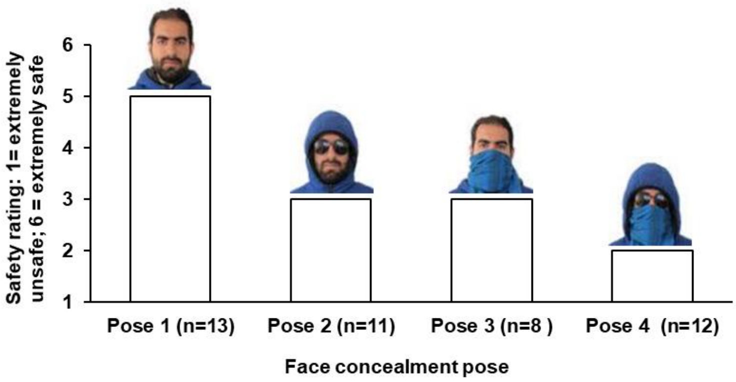


Figure 5.8. Median rating across all actors for the first evaluations in the face block (has order bias).

The second analysis excluded images used in the hand and/or mixed blocks. In detail, the calculation used the median evaluations of the first trial in the face block when this was the first test; this reduced the data sample from n = 44 (44 x 1) to n = 16 (16 x 1). Figure 5.9 presents these responses, which are almost identical to the whole set in Figure 5.2. The only difference is that the median of Pose 2 (top of the face concealed) increased from 3.00 to 3.50. However, it was still not completely free of order bias, as the evaluations may have been influenced by those of the paired comparison procedure. This means that all the images in this analysis were the first evaluated image in the face block, but some images were rated first in the paired comparison procedure. The next analysis examines this effect.

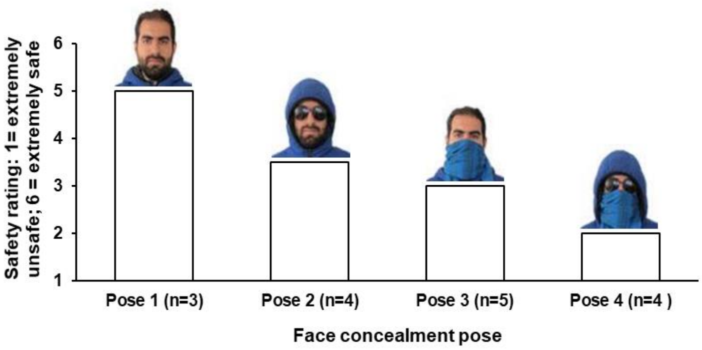


Figure 5.9. Median rating across all actors for the first evaluations in the face block (has least order bias).

The third analysis excluded images used in the paired comparison procedure, giving only the first images in the rating data. This means that all the images were the first to be evaluated in the face block, and had not previously been seen. Thus, participants did not have a clear idea as to what the experiment was about, and so these evaluations were completely free of order bias.

The analysis calculated the median evaluations for the face targets with no order bias. This reduced the data sample from n = 44 (44 x 1) to n = 11 (16 x 1). Figure 5.10 presents the responses for each face pose of this sample, the results of which differ from the whole data set (Figure 5.2), and from the sample with the least order bias (Figure 5.9). The major difference is that the median for perceived safety in Pose 4 (face concealed) increased from 2 to 4. It is important to note that this median (4) involved only two ratings from male participants, and the gender target on the two images was one male in the image and one female in the other image.

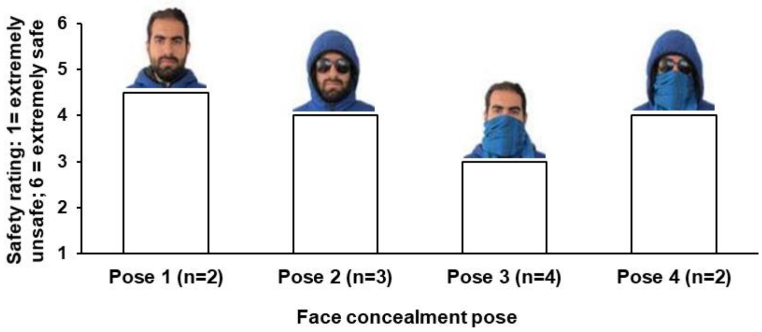


Figure 5.10. Median rating across all actors for the first evaluations in the face block (free of order bias).

In summary, the first responses were analysed to test the order effect and assess the impact of previous evaluations. This analysis used only the face block data. If the first rating was the same as the final overall rating, the results were not biased by previous evaluation/s. Due to the randomisation effect, a small sample size (n = 16) was tested with least order bias. Figure 5.11 compares the overall rating and the first responses (with least order bias) for the four face poses. The only difference is that the median rating for perceived safety in Pose 2 (top of the face concealed) increased from 3.00 to 3.50. Responses did not change much between the images that rated first and that were not rated first. This suggests these data were not likely to have been biased by previous evaluations. Thus, the identical images that rated in different blocks were evaluated similarly.

One limitation of this approach is that the sample size was reduced due to the randomised target order within a block and procedure, meaning that the number of ratings per individual target ranged from 0 to 3. For each face pose, there were 2 to 13 individual evaluations.

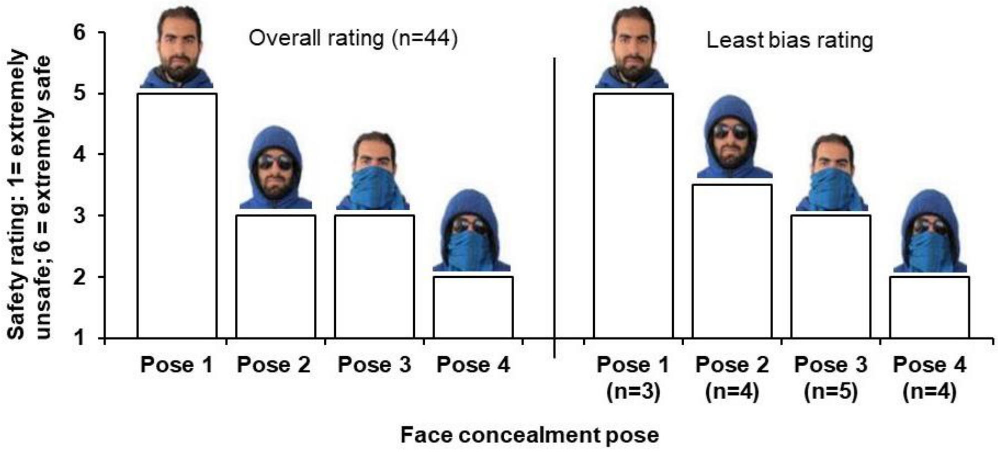


Figure 5.11. Comparison between overall ratings (whole data set) and first response ratings (least order bias) for the four face poses in the face block.

5.2.3. Analysis of using two different rating scales

For Experiment 2, a 6-point rating scale (1 = extremely unsafe, 2 = very unsafe, 3 = slightly unsafe, 4 = slightly safe, 5 = very safe, 6 = extremely safe) was used as the main rating scale in the three test blocks. At the end of the experiment and after collecting data from all blocks, participants were asked to repeat the face block using a 5-point rating scale (1 = very unsafe, 2 = somewhat unsafe, 3= neutral , 4 = somewhat safe , 5 = very safe) used in Experiment 1. These data were collected from the whole sample (n = 44 x 16 images) for the face block. The aim was to compare evaluations using 5-point and 6-point response scales and to see whether the responses differed (see Section 3.4.1). This analysis addressed two questions:

Q1. Is the middle choice (e.g., 3 = neutral) a common and easy choice when available?

Q2. Does the use of two different scales change the results?

To answer Q1, a brief analysis was conducted on the responses and revealed that the number of times the middle choice was selected was 132/704 (18.75%), and four of 32 participants never chose the middle option. This suggested that the middle option (3 = neutral) could be an easy choice. Nevertheless, further in-depth analysis (Q2) was required to confirm this.

For Q2, the results from the two scales (5-points and 6-points) were normalised into a new 20-point scale. Several mathematical steps were applied to change the two scales. First, one value was subtracted from the response values for each scale, and the result was then divided by the maximum choice: 5 for the 5-point scale and 6 for the 6-point scale, as formulated in the equation below.

$$\frac{Response\ -1}{Scale\ type\ (5,6)} \times New\ Scale\ (20)$$

This equation ensured the maximum and minimum value on the original scales were the maximum and minimum values on the new scale. For example, if the image rated the minimum (1) in the original scales, it was the minimum (1) in the new scale; and if it rated the maximum (5) in the 5-point scale or (6) in the 6-point scale, it was the maximum (20) in the new scale. The rest of the values were then equally distributed in the middle.

Figure 5.12 presents the converted responses from 6-point and 5-point scales to the new 20-point scale; it depicts the median rating across the four actors with the same pose in the face block. On the new 20-point scale, the evaluations of face poses between the 6-point and 5-point scales differs. For example, Pose 1 rated 16.7 on the 6-point scale, while on the 5-point scale it rated 8 – a large difference of 8.7 points. Similarly, Poses 2 and 3 differ by 2 points, and Pose 4 differs by 1.3 points.

Wilcoxon test was used to determine the significance level. This revealed a significant difference (p = 0.013) between responses using the 5-point scale and responses using the 6-point scale. This indicated different evaluations between the two scales. For greater accuracy, the next analysis investigated the consistency between the responses of the two different scales.

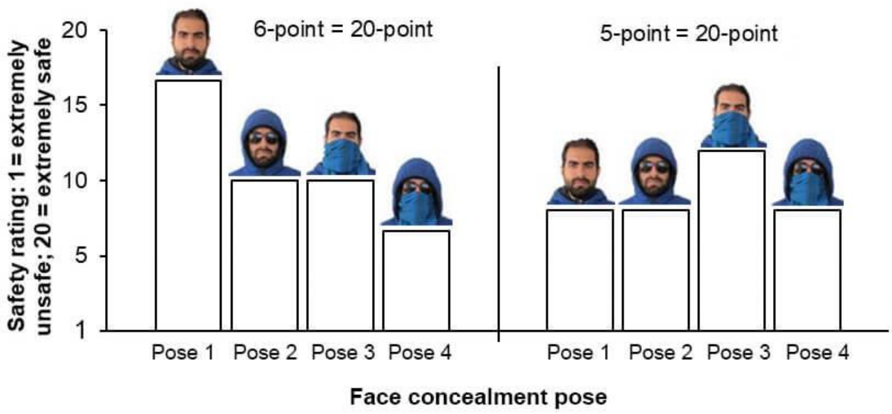


Figure 5.12. Comparing the safety rating for the four face poses between two different rating scales (n = 44). *Left*: converted responses from the 5-point scale to the 20-point scale. *Right:* converted responses from the 6-point scale to the 20-point scale.

There were some responses that were strangely rated differently than most of the others. These were identified as outliers as the scores differed substantially from the remainder of the data (Field, 2009, p. 98). The analysis identified participants who had outliers in their ratings of the 16 images using the boxplot graph generated in the normality test. Among the 44 participants, a total of 12/704 (1.7%) outliers were found.

The distribution of the 12 outliers among the 44 participants revealed a higher number (3) for participant ID 26; this participant rated 3/16 (19%) images very differently from most ratings. The rest of the nine outliers were distributed equally between nine participants with one outlier for each (6%). The distribution of the 12 outliers among the 16 images was as follows: 4 for Image 12; 2 for Images 6, 13, and 14; and 1 for Images 9 and 15. Having this very small number (12/704) of outlier responses does not seem to make a difference, and the results did not change after removing them (Figure 5.13: Original and ≥50%). Therefore, the outliers were retained in the analysis.

Further analysis was conducted by setting different thresholds of inconsistent responses to see how this affected the results. This was done by excluding the inconsistent responses on both converted scales (5 and 6 points). Thus, when a difference between the responses of the same images reached a certain threshold (≥ 50%, ≥ 40%, ≥ 35%), the responses of this participant were excluded from the analysis. Using this approach, Figures 5.13 and 5.14 present the safety responses on the 6-point and 5-point scales converted to the new 20-point scale, respectively. In both scales, for the threshold ≥ 50%

and ≥ 40%, no difference was found; for the ≥ 35%, responses were slightly different. Overall, no significant differences were found after removing inconsistent answers.

In conclusion, the new standardised graphs of the responses on the 5-point and 6-point scales of the same image appear different in shape (Figure 5.14). Indeed, the Wilcoxon test revealed a significant difference (p = 0.015), indicating that using a different scale type could change the results. However, this conclusion has an order bias as participants completed the 5-point scale at the end of the experiment, approximately 45 minutes (including the time spent conducting the paired comparison procedure) after evaluating images using the 6-point scale.

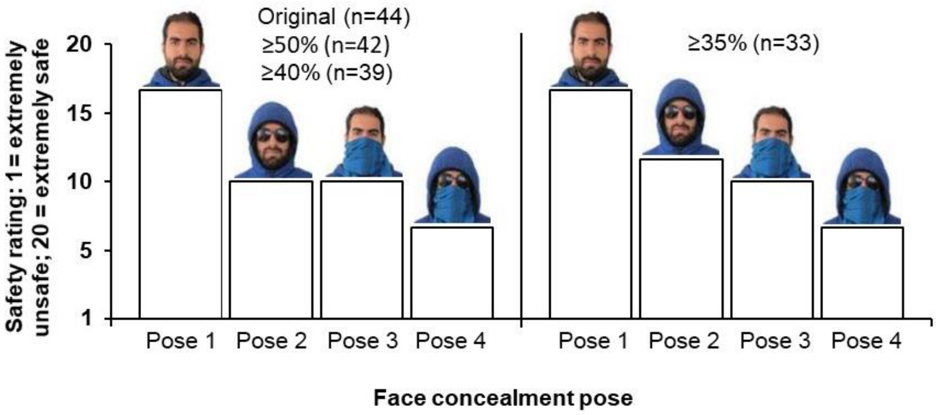


Figure 5.13. Different thresholds for excluding inconsistent responses on the *6-point scale* converted to the new 20-point scale.

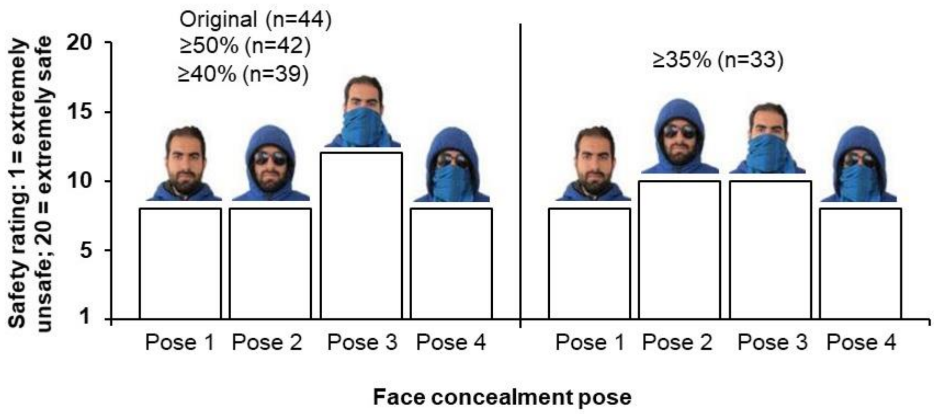


Figure 5.14. Different thresholds for excluding inconsistent responses on the *5-point scale* converted to the new 20-point scale.

5.3. Results: Paired comparison

In the paired comparison procedure in Experiment 2, the null condition trials were those where identical images were shown on the left-hand and right-hand sides of the pair. In each block of trials, there were eight null-condition pairs, chosen randomly from the 16 test images available in a block, labelled here as images 1 to 8 (Tables 5.9 to 5.12). These eight null condition pairs were presented, in a random order, within the main block of 120 trials. Each of the 44 participants provided 24 null-condition responses (8 nulls × 3 blocks), these responses were then inspected to determine tendencies towards position bias.

Tables 5.9 to 5.12 presents the percentages of votes given to the right-hand image of each null image in each scene and block, and across all blocks in a scene. In fact, within each scene and block, the right-hand image was considered safer in a range from 52% to 84%. Across both scenes, the right-hand image was considered safer in 63% of null condition trials in the face block, 66% of the hands block, 71% of the mixed block, and 67% across all blocks, suggesting a tendency to indicate the right-hand image more frequently than the left-hand image.

Table 5.9. Number and percentage of votes for the right-hand image for each null image trial in Scene 1 Scene 2, and both scenes together for the face block.

Null image no.	**Scene 1**		**Scene 2**		**Scene 1&2**	
	(n = 22)	**%**	**(n = 22)**	**%**	**(n = 44)**	**%**
1	13	59%	14	64%	27	61%
2	11	50%	12	55%	23	52%
3	12	55%	16	73%	28	64%
4	18	82%	15	68%	33	75%
5	14	64%	11	50%	25	57%
6	15	68%	14	64%	29	66%
7	16	73%	17	77%	33	75%
8	12	55%	13	59%	25	57%
Total	111	63%	112	64%	223	63%

If position bias was not a significant factor, then the left and right images would each be identified as safer in 50% of the null condition trials. A Z-test (for two independent proportions) suggested that the bias to the right-hand image in null condition trials was significant ($z = 9.02$, $p < 0.05$) (Table 5.13).

Table 5.10. Number and percentage of votes for the right-hand image for each null image trial in Scene 1 Scene 2, and both scenes together for the hands block.

Null image no.	Scene 1		Scene 2		Scene 1&2	
	(n = 22)	%	(n = 22)	%	(n = 44)	%
1	16	73%	15	68%	31	70%
2	13	59%	12	55%	25	57%
3	15	68%	17	77%	32	73%
4	15	68%	15	68%	30	68%
5	12	55%	14	64%	26	59%
6	12	55%	16	73%	28	64%
7	15	68%	14	64%	29	66%
8	16	73%	16	73%	32	73%
Total	114	65%	119	68%	233	66%

Table 5.11. Number and percentage of votes for the right-hand image for each null image trial in Scene 1 Scene 2, and both scenes together for the mixed block.

Null image no.	Scene 1		Scene 2		Scene 1&2	
	(n = 22)	%	(n = 22)	%	(n = 44)	%
1	15	68%	15	68%	30	68%
2	14	64%	16	73%	30	68%
3	15	68%	14	64%	29	66%
4	14	64%	16	73%	30	68%
5	19	86%	14	64%	33	75%
6	14	64%	16	73%	30	68%
7	14	64%	17	77%	31	70%
8	19	86%	18	82%	37	84%
Total	124	70%	126	72%	250	71%

Table 5.12. Number and percentage of votes for the right-hand image for each null image trial in Scene 1 Scene 2, and both scenes together across all blocks.

Null image no.	Scene 1		Scene 2		Scene 1&2	
	(n = 22)	%	(n = 22)	%	(n = 44)	%
1	44	67%	44	67%	88	67%
2	38	58%	40	61%	78	59%
3	42	64%	47	71%	89	67%
4	47	71%	46	70%	93	70%
5	45	68%	39	59%	84	64%
6	41	62%	46	70%	87	66%
7	45	68%	48	73%	93	70%
8	47	71%	47	71%	94	71%
Total	349	66%	357	68%	706	67%

Table 5.13. Results for the Z test of position bias in the null-condition trials of images for each scene and block, and across all blocks.

Block	Scene	n.	Number null pairs	Target if no bias	Percentage right	Number of cases	Z based on cases*	Sig. bias
Face	S1	22	8	0.5	0.63	176	3.45	Yes
	S2	22	8	0.5	0.64	176	3.71	Yes
Hand	S1	22	8	0.5	0.65	176	3.98	Yes
	S2	22	8	0.5	0.68	176	4.78	Yes
Mixed	S1	22	8	0.5	0.70	176	5.31	Yes
	S2	22	8	0.5	0.72	176	5.84	Yes
Overall	S1	22	24	0.5	0.66	528	7.35	Yes
	S2	22	24	0.5	0.68	528	8.27	Yes
	S1&2	44	24	0.5	0.67	1056	11.05	Yes

* Significant at the 0.05 level when Z is less than -1.96 or greater than 1.96 (Bland, 2000).

The next question to address is whether this position bias persisted within the main images. Figure 5.15 plots data from all blocks and scenes, and presents percentage votes for the right-hand images in the main data (Y axis) against the null condition trials (X axis). This does not suggest position bias persisted beyond null condition trials. There is no correlation between the null condition trials and the main trials where it was centralised around 50%, which denotes a balance between right and left side responses.

In summary, while the null condition trials revealed a significant bias towards the right-hand image, this trend was not revealed between-images within the main data. In the main data, the right-hand images were considered safer in 49% of main trials in the face block and 51% of the hands block; and the Z-test does not suggest a significant bias ($z = -34.25$, $p < 0.05$) to the right-hand image in main data trials. In any case, image position (left or right) was randomised to offset position bias (Veitch et al., 2019). Therefore, this bias towards the right-hand image was equally distributed among both images in a pair and should not have significantly affected the results.

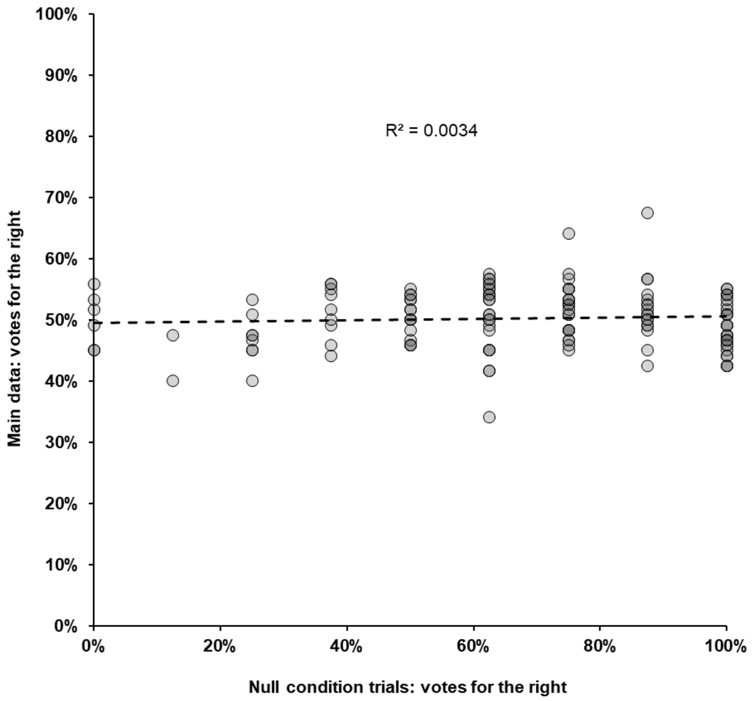


Figure 5.15. Percentage of votes for the right-hand image of the null condition trials (X axis) against percentage of votes for the right-hand image of the main data (Y axis).

Two different background scenes were used in the test images (see Figure 3.2) to determine whether changes in the background scene led to different evaluations. Each scene used in trials was evaluated by a different sample of 22 test participants. Each participant evaluated images with a given background scene, giving a maximum of 330 (15 pairs x 22 participants) votes per scene and block. Table 5.14 presents the number and percentage of votes for each test image in each block and scene along with the statistical test applied. For the face block, the percentages of votes for the same image in Scene 1 and Scene 2 were very similar; no differences greater than 10% were found between the two scenes. For the hands block, the differences were less than 10% except for Images 7 (12%) and 14 (17%). For the mixed block, no differences greater than 10% were found between the two scenes. These observations suggest that differences between the scenes were negligible.

Different samples were used in trials with different background scenes; to test for significant differences, the Mann-Whitney test for independent samples was used (Table 5.14). For the face block and the mixed block, there were no significant differences between background scenes. For the hands block, there were significant differences ($p < 0.05$) for three of the 16 images (no. 4, 7, and 14). In further analyses, the data were reorganised as one set of 44 test participants, and the background scene was ignored.

Table 5.14. Votes allocated to scenes as a safer image for the two background scenes and the *p*-values** derived using Mann-Whitney for each image in each block (Maximum vote 330).

Image no.	Face block						Hands block						Mixed block					
	Scene 1		Scene 2		Dif.*	Mann-Whitney test	Scene 1		Scene 2		Dif.*	Mann-Whitney test	Scene 1		Scene 2		Dif.*	Mann-Whitney test
	n	%	n	%			n	%	n	%			n	%	n	%		
1	235	71%	242	73%	2%	p=0.548	153	46%	156	47%	1%	p=0.804	239	72%	233	71%	1%	p=0.943
2	122	37%	108	33%	4%	p=0.448	116	35%	141	43%	8%	p=0.174	73	22%	83	25%	3%	p=0.476
3	124	38%	126	38%	0%	p=0.934	112	34%	95	29%	5%	p=0.165	176	53%	164	50%	3%	p=0.623
4	17	5%	6	2%	3%	p=0.028	39	12%	14	4%	8%	p=0.032	22	7%	19	6%	1%	p=0.744
5	265	80%	243	74%	6%	p=0.098	200	61%	190	58%	3%	p=0.768	265	80%	246	75%	5%	p=0.225
6	133	40%	113	34%	6%	p=0.165	140	42%	164	50%	8%	p=0.266	89	27%	94	28%	1%	p=0.575
7	134	41%	145	44%	3%	p=0.594	164	50%	126	38%	12%	p=0.022	188	57%	189	57%	0%	p=0.701
8	23	7%	29	9%	2%	p=0.130	64	19%	34	10%	9%	p=0.567	38	12%	34	10%	2%	p=0.884
9	316	96%	317	96%	0%	p=0.800	276	84%	280	85%	1%	p=0.971	306	93%	309	94%	1%	p=0.990
10	222	67%	212	64%	3%	p=0.578	230	70%	255	77%	7%	p=0.261	158	48%	149	45%	3%	p=0.239
11	216	65%	230	70%	5%	p=0.335	226	68%	205	62%	6%	p=0.292	259	78%	258	78%	0%	p=0.617
12	94	28%	90	27%	1%	p=0.728	160	48%	126	38%	10%	p=0.224	91	28%	84	25%	3%	p=0.678
13	295	89%	309	94%	5%	p=0.637	247	75%	276	84%	9%	p=0.154	293	89%	307	93%	4%	p=0.275
14	187	57%	192	58%	1%	p=0.536	205	62%	260	79%	17%	p=0.030	146	44%	149	45%	1%	p=0.613
15	190	58%	204	62%	4%	p=0.451	206	62%	202	61%	1%	p=0.943	215	65%	236	72%	7%	p=0.174
16	67	20%	74	22%	2%	p=0.392	102	31%	116	35%	4%	p=0.281	82	25%	86	26%	1%	p=0.755

* The difference between the percentage of votes for Scenes 1 and 2.
** $p < 0.05$ indicates a significant effect.

5.3.1. Analysis of the test blocks

In each block of paired comparison procedure, the 16 test images (4 actors x 4 pose variations: see Table 3.8) were evaluated in all possible combinations of 120 pairs (excluding null conditions). For each block, an evaluation was required to indicate the safer situation of the two images. Therefore, a lower number of votes meant there were fewer occasions on which an image was considered the safer of the pair.

To identify the significance Differences between poses, the scaling method of Dunn-Rankin (VSRS) was used where the difference between item sums ≥ 37.1 indicates a significant difference ($p < 0.01$) (see Section 3.4.2).

In the following analyses where the two background scenes were combined, for the 16 test images, the maximum possible number of votes for each was 660 votes (each paired with 15 other images × 44 participants), indicating the percentage of evaluations in which the pose was stated to be the safer of the two. A lower percentage indicates a pose was considered to be less safe. For the face block, the percentages of votes for an image being considered safer are presented in Figure 5.16, which shows the distribution of votes as safer for each pose of an actor. In general, a person was considered safer when the face was exposed, and females were considered safer than males. Poses with either the bottom or the top of face concealed exhibited similar trends. The images awarded the greatest and lowest numbers of votes are depicted in Figure 5.17.

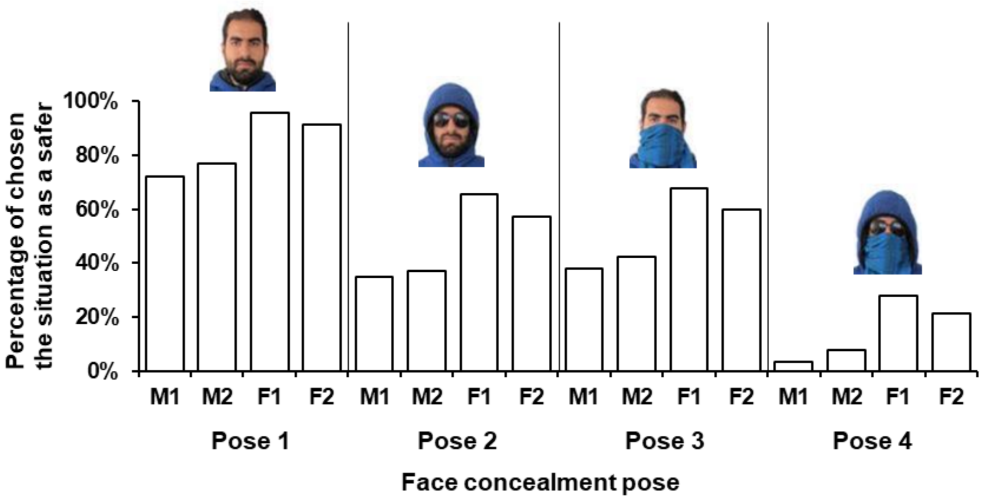


Figure 5.16. Percentage of votes as a safer situation for each pose of an actor in the face block. Information about each actor were illustrated in Table 3.3.

Figure 5.17. Images awarded the greatest and fewest numbers of votes in the face block. *(a)* The least safe situation. *(b)* The most safe situation.

The percentage of votes across the four actors with the same face pose are presented in Figure 5.18. Face concealment reveals a gradual reduction in votes as the safer situation as the face changes from fully exposed (84%) to fully concealed (15%). In the middle, the percentage of votes for top face concealed (49%) and bottom face concealed (52%) were close with a difference of 3%. This conclusion was consistent for the four actors (Figure 5.16). This indicates that as face exposure was reduced by wearing hood and sunglasses and/or scarf, the feeling of safety was also reduced.

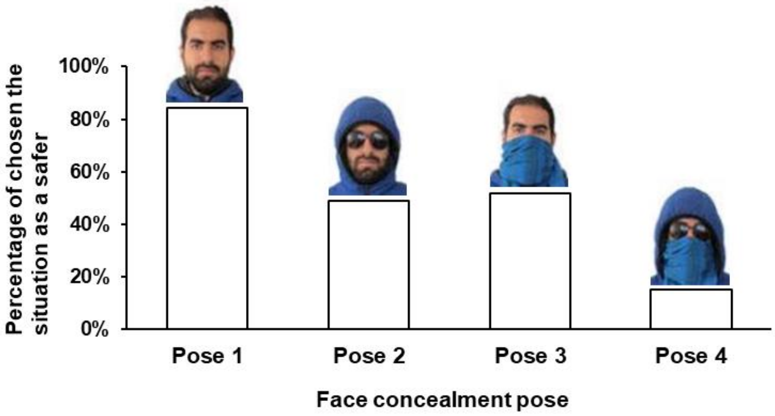


Figure 5.18. Percentage of votes as a safer situation across all actors for each pose in the face block.

The Dunn-Rankin analysis used scale values to show the preference between the pair images; which are the proportional to the sum of the ranks assigned by the choices to each of the poses (Dunn-Rankin et al., 2014, p. 55), and are obtained by dividing each vote total by the maximum vote possible and multiplying by 100 (Dunn-Rankin et al., 2014, p. 56).

Figure 5.19 presents the rank between the face poses in Dunn-Rankin scale values across all the actors. It illustrates that changing the exposure of the face led to different evaluations. The exposed face (Pose 1) was ranked highest in terms of being safer at 92.7/96; the whole face concealed (Pose 4) was ranked lowest at 6.4/96, whereas a similar value was generated when the top of the face was concealed (Pose 3) at 48.5/96, and when the bottom of the face was concealed (Pose 3) at 52.3/96. Table 5.15 provides scale value and sums for each pose.

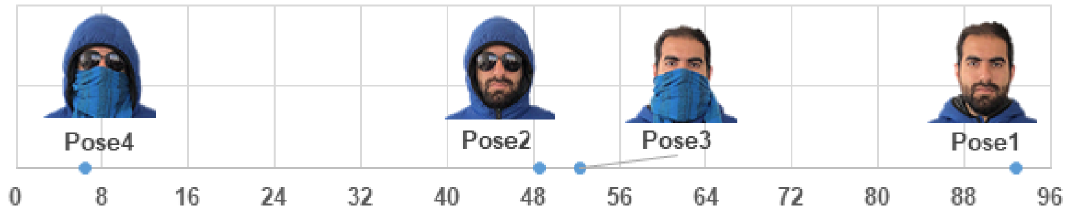


Figure 5.19. Dunn-Rankin scale value across all the actors in the face block, where a higher number denotes a safer rating.

Table 5.15. The scale value and sum for each pose in the face block. The poses were illustrated in Figure 5.19.

Face pose	**Scale Value/96**	**Sum/132**
Pose 1	92.7	122.4
Pose 2	48.5	64.1
Pose 3	52.3	69.1
Pose 4	6.4	8.5

The significance of differences between poses is shown in Table 5.16. The Dunn-Rankin (VSRS) revealed no significant differences between top concealed faces (Pose 3) and bottom concealed faces (Pose 2), but significant differences ($p < 0.01$) between the other poses (1 vs 2, 1 vs 3, 1 vs 4, 2 vs 4, 3 vs 4). In summary, Pose 1; where the face was exposed, was perceived to be significantly safer than the Poses 2, 3 or 4 where the face was at least partially concealed by clothing.

Table 5.16. Significance of differences between poses in the face block, as determined using Dunn-Rankin (VSRS); the significant values can be found in Appendix K, Table K.1.

Face part obscured	**Top** (Pose 2)	**Bottom** (Pose 3)	**All** (Pose 4)
None (Pose 1)	p<0.01	p<0.01	p<0.01
Top (Pose 2)	---	n.s.	p<0.01
Bottom (Pose 3)	---	---	p<0.01

For the hands block, the percentages of votes for an image being considered safer are presented in Figure 5.20, which shows the distribution of votes for each pose of an actor as safer. Generally, a person was considered less safe when their hands were located behind their body, and females were considered safer than males. The three poses: hands at the sides of the body, in the front of the body, and in the pocket exhibited similar trends. The images awarded the greatest and fewest numbers of votes are depicted in Figure 5.21.

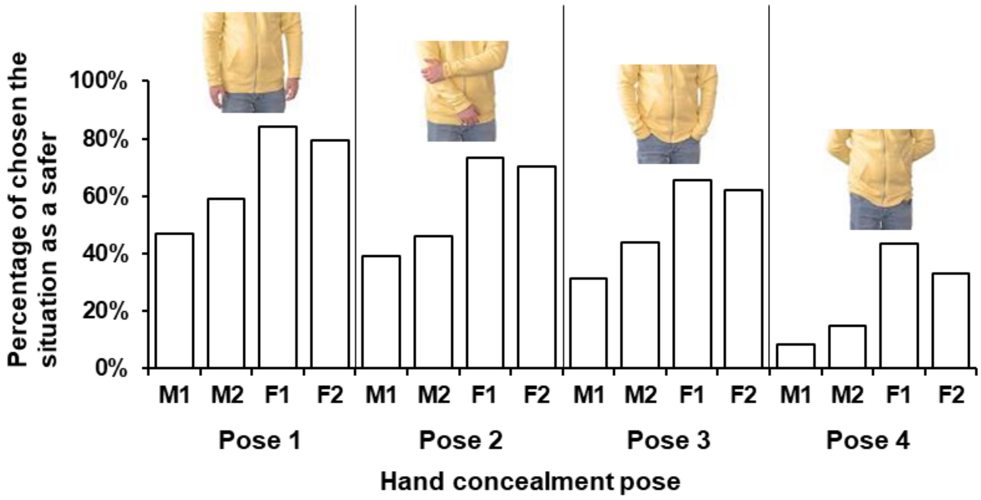


Figure 5.20. Percentage of votes as a safer situation for each pose of an actor in the hands block. Information about each actor were illustrated in Table 3.3.

Figure 5.21. Images awarded the greatest and fewest numbers of votes in the hands block. *(a)* The least safe situation. *(b)* The most safe situation.

The percentage of votes across the four actors with the same hands pose are depicted in Figure 5.22. There was a slight and gradual reduction in safety votes as the hands changed from being fully exposed (67%) to fully concealed (25%); when hand exposure was reduced by placing the hands in the pockets or behind the body, the feeling of safety was reduced. Hands in the front of the body and inside the pocket received similar numbers of votes, with a difference between them of only 6%. This conclusion was consistent for the four actors (Figure 5.20). This indicates that when hands are concealed by putting them in pockets or behind the body, the feeling of safety is reduced.

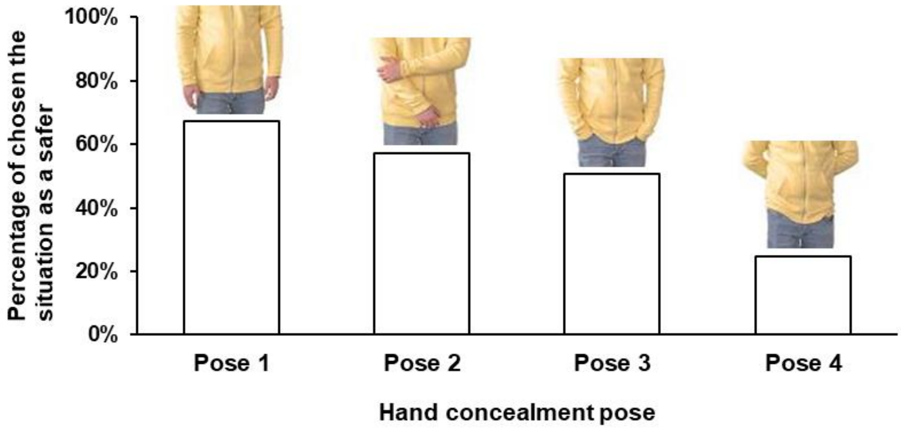


Figure 5.22. Percentage of votes as a safer situation across all actors for each pose in the hands block.

Figure 5.23 presents the rank between the hand poses in Dunn-Rankin scale values across all the actors. This shows that changing exposure of the hands led to different evaluations; the exposed hands at the sides (Pose 1) and hands in front (Pose 2) were ranked highest in terms of safety (92.7/96 and 59.0/96, respectively). The lowest rank was given when the hands were concealed behind (Pose 4) at 18.5/96, while a somewhat different value was given when the hands were concealed in pockets (Pose 3) at 50.8/96. Table 5.17 provides scale value and sums for each pose.

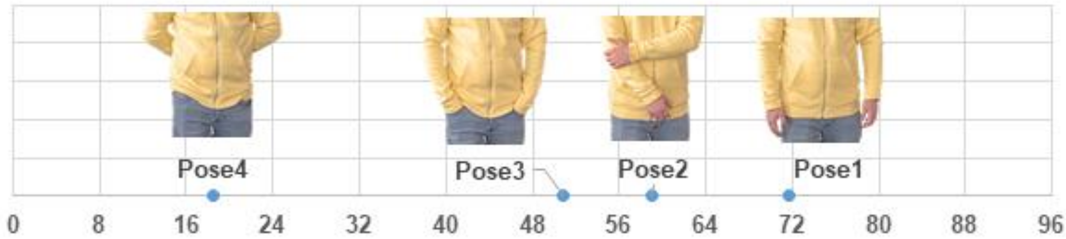


Figure 5.23. Dunn-Rankin scale value across all the actors in the hands block, where a higher number denotes a safer rating.

Table 5.17. The scale value and sum for each pose in the hands block. The poses are illustrated in Figure 5.23.

Hands position	Scale Value/96	Sum/132
Side (Pose 1)	71.7	94.6
Front (Pose 2)	59.0	77.9
Pocket (Pose 3)	50.8	67.0
Behind (Pose 4)	18.5	24.4

The significance of the differences between hand poses are displayed in Table 5.18. The Dunn-Rankin (VSRS) identified a significant difference ($p < 0.01$) between hands behind (Pose 4) and all other hand positions (Poses 1, 2, and 3). The differences between the latter were not found to be significant. In summary, changing the exposure of hands by putting them in different positions did not exert a large impact on the feeling of safety; and concealing the hands in the pockets had less effect on safety, while concealing the hands behind had a greater effect. This suggested that hands behind the body considered the less safe position.

Table 5.18. Significance of differences between poses in the hands block, as determined using Dunn-Rankin (VSRS); the significant values are shown in Appendix K, Table K.2.

Hands position	**Front** (Pose 2)	**Pocket** (Pose 3)	**Behind** (Pose 4)
Side (Pose 1)	n.s.	n.s.	p<0.01
Front (Pose 2)	---	n.s.	p<0.01
Pocket (Pose 3)	---	---	p<0.01

For the mixed block, the percentages of votes for an image being considered safer are presented in Figure 5.24 which shows the distribution of votes as safer for each pose of an actor. In general, a person was considered less safe if the face and the hands were concealed, and females were considered safer than males. Face exposed and hands concealed (Pose 3) exhibited to be safer than the converse pose – face concealed and hands exposed (Pose 4). The mixed conditions of the highest and lowest votes as safer are depicted in Figure 5.25.

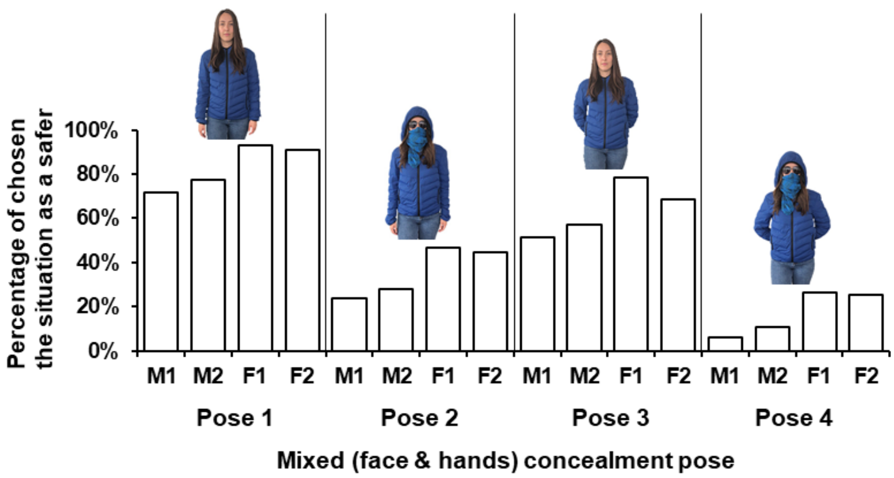


Figure 5.24. Percentage of votes as a safer situation for each pose of an actor in the mixed block. Information about each actor were illustrated in Table 3.3.

Figure 5.25. Images awarded the greatest and fewest numbers of votes in the mixed block *(a)* The least safe situation. *(b)* The most safe situation.

The percentage of votes across the four actors in the same mixed condition are presented in Figure 5.26. There was a large difference (66%) between fully exposed face and hands (Pose 1) and fully concealed face and hands (Pose 4). Also, when the face was exposed (Poses 1 and 3), the safety rating was higher than when the face was concealed (Poses 2 and 4). This indicates the importance of exposing the face and hands, but begs a question, which is more important? As an answer, the difference in votes regarding the effect of face covering was greater (47%) than the difference in votes regarding the effect of hand covering (19%). This conclusion was the same whether hands were exposed or concealed (face effect), and whether the face was exposed or concealed (hands effect). This conclusion was consistent for the four actors (Figure 5.24). This indicates that face exposure could be more important than hand exposure.

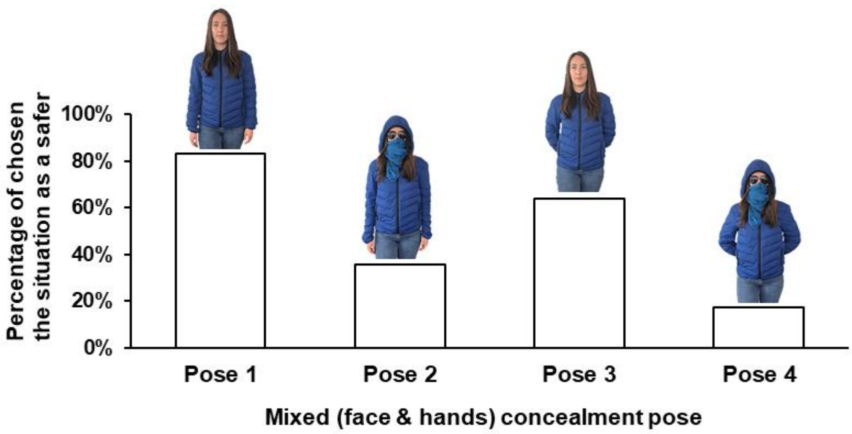


Figure 5.26. Percentage of votes as a safer situation across all actors for each pose in the mixed block.

Figure 5.27 presents the rank between the mixed poses in Dunn-Rankin scale values across all the actors. This shows different evaluations for the mixed conditions; the full exposure of face and hands (Pose 1) was ranked highest at 91.6/96, while full concealment of face and hands (Pose 4) was ranked the lowest at 9.1/96. In between, face concealed and hands exposed (Pose 2) ranked lowest at 32.1/96 than the converse pose – face exposed and hands concealed (Pose 3) at 67.3/96. Table 5.19 provides scale value and sums for each pose.

The comparison between Poses 1 and 2 indicates that face exposure has an influence the feeling of safety which is the same conclusion was reached for the face block (Figure 5.18). Furthermore, the comparison between Poses 1 and 3 indicates that hand exposure has an influence the feeling of safety which is the same conclusion was reached for the hands block (Figure 5.22).

Figure 5.27. Dunn-Rankin scale value across all the actors in the mixed block, where a higher number denotes a safer rating.

Table 5.19. The scale value and sum for each pose in the mixed block. The poses were illustrated in Figure 5.27.

Mixed condition	Scale Value/96	Sum/132
Face and hands exposed (Pose 1)	91.6	120.9
Face concealed, hands exposed (Pose 2)	32.1	42.3
Face exposed, hands concealed (Pose 3)	67.3	88.8
Face and hands concealed (Pose 4)	9.1	12.0

The significant differences between the mixed poses are presented in Table 5.29. Dunn-Rankin (VSRS) revealed a significant difference ($p < 0.01$) between all the mixed conditions, except between face concealed and hands exposed (Pose 2) and face concealed and hands concealed (Pose 4). In this block, two comparisons regarding the hands effect can be made, which is between Poses 1 v 3 and 2 v 4; VSRS suggested one significant difference (1 v 3) and one is not (1 v 3). This could lead to the same conclusion of the hands block where the hand concealments did not have a big effect on feelings of safety. Likewise, two comparisons regarding the face effect can be made, which is between Poses 1 v 2 and 3 v 4; VSRS suggested a significant difference in both comparisons. This leads to the same conclusion of the face block where the face concealments have a big effect on feelings of safety.

Table 5.20. Significance of differences between poses in the mixed block, as determined using Dunn-Rankin (VSRS); the significant values are presented in Appendix K, Table K.3.

Mixed condition	**Face concealed hands exposed** (pose 2)	**Face exposed, hands concealed** (pose 3)	**Face and hands concealed** (pose 4)
Face and hands exposed (Pose 1)	$p<0.01$	$p<0.01$	$p<0.01$
Face concealed, hands exposed (Pose 2)	---	$p<0.01$	n.s.
Face exposed, hands concealed (Pose 3)	---	---	$p<0.01$

5.4. Summary

Experiment 2 compared the importance of the face and the hands by varying the degree to which these features were concealed using two procedures: category rating and paired comparisons (described in Sections 3.4.1 and 3.4.2, respectively). Each procedure involved evaluating 16 test images in three blocks: face, hands, and mixed. The experiment evaluated the same image using two different background scenes but found no significant differences between them; both procedures and all blocks reached the same conclusion.

In the category rating procedure, the order bias analysis using the repeated images did not suggest an order bias on the evaluations. Analysis of identical images rated in different blocks did not suggest a rang bias, which means no effect of a previous image evaluation was found. Results from the three test blocks suggested that exposure of the face was important for evaluating other pedestrians because the changes of face concealment led to different evaluations of safety, when face concealment increased, the safety rating decreased (a less safe feeling). Moreover, being able to see the hands had little effect, if any; it was a smaller effect than found with the face. Analysis of using two different rating scales applied for face block suggested that changing the scale type might change the results; nevertheless, there was an order bias effect on this conclusion.

In the paired comparison procedure, the null condition data suggested a significant and consistent tendency to choose the right-hand image of the pair. This position bias was not, however, revealed in the main images trials, and in any case image position was randomised between subjects to offset this effect; this conclusion was the same for all blocks. Results from the face block indicated that covering the whole face affected feelings of safety. From the hands block, concealing the hands slightly affected feelings of safety, but the position of hands behind the body was shown to exert a greater effect than the other positions tested (at the sides of the body, in the front of the body, and in the pocket). The results from the mixed block supported those for the face and hands blocks, where seeing the face is important in perceived safety, and seeing the hands is less important.

Regarding the effect of gender, in each test block for both procedures, female actors were evaluated as safer than the male actors, confirming the findings of Experiment 1 (see Section 4.3.1). This difference was expected because males are considered to be more fearless and fear-provoking than females (Koskela, 1997; Pain, 2001).

Experiments 1 and 2 investigated a range of potential visual cues employed in interpersonal evaluations of pedestrians. The results suggested that the face was an important visual cue. Nevertheless, that finding could be a result of the methodology employed; for example, unintentionally leading respondents to favour the face in subjective evaluations. Therefore, to verify these findings, Experiment 3 used an entirely different approach, eye-tracking. The eye-tracking provides an objective measure of visual attention, by being an alternative procedure for subjective evaluation, and therefore offers an independent approach to test the veracity of the conclusion that the face is the key visual cue when looking at other people.

Chapter 6. Eye-Tracking – Experiment 3: Method

6.1. Introduction

This chapter describes the method employed in Experiment 3 which was conducted to investigate visual cues using an alternative procedure to the previous two experiments. In this experiment, an eye-tracking device was used to measure gaze behaviour which in a given context indicates how people interact visually in an environment, which can reveal important differences between the visual cues in the observed scene.

Looking at something is called a “visual fixation”, and is defined as a temporary stopping of the eye at a certain place, which can last from milliseconds up to several seconds (Holmqvist et al., 2011, p.51). Fixating on something means it has the individual’s cognitive attention and is related to any subsequent action. For instance, before washing hands, people fixate on the tap, soap, and paper towels, then fixate on the tap to wash their hands (Hayhoe et al., 2003). Eye-tracking glasses provide detailed information about eye movements such as times and duration of fixations, saccades, blinks, and the coordinates of the gaze position. This experiment measured the duration of all fixations on the different domains (e.g. face, vehicle) of a scene along with the location of first fixation within each trial. It was designed to test the following hypotheses:

1. Visual fixation (order and duration) will be greater on the person than the surroundings.
2. Visual fixation (order and duration) will be greater on the face than on the body.

Experiment 3 repeated the method employed by Fletcher-Watson et al. (2008) who used gaze distribution in naturalistic scenes to investigate whether a person would be gazed at without it being a search task. They used side-by-side images, one of which contained a person, and one of which did not (see Figure 2.7). These were observed in two blocks of trials: free viewing with no other task, and viewing with the task of discriminating the gender of the person. Eye-tracking glasses were used to measure the time spent looking at each domain of the scene and the location of the first fixation within each trial. In the analysis, the areas were weighted to give relative measures between domains because faces (for example) are always considerably smaller than bodies and backgrounds. Their results suggested a strong tendency to gaze towards the person (specifically, directly to the face).

One limitation of Fletcher-Watson et al.'s (2008) study is that they did not build on the validation steps recommended by Veitch et al. (2019). For example, it was assumed that there was an involuntary bias towards the face where the person size in the pair image was large (almost 50% of total image size) which made it easy to gaze toward it (See Figure 2.7); nor did it include validation steps (control conditions); for example, distraction factors such as animals, vehicles, and/or signs. Therefore, Experiment 3 extended Fletcher-Watson et al.'s study by including control variables to confirm that when people looked at the targets' faces, these were what they intended to look at and not something else.

6.2. Test images

Participants were asked to observe a series of images projected onto a screen while wearing eye-tracking glasses. Test images were presented using a projector (Optoma HD31UST) to cast the images onto a white wall with dimensions 1559 × 875 PPI at a viewable size of 2.21 x 1.25 m (same setting as Experiment 2: see Section 3.2).

Experiment 3 aimed to determine the likelihood of looking at the person rather than something else (e.g., the background), and if so, whether the tendency was to look predominantly at the face. Therefore, different versions of test images were created using various target persons and different backgrounds.

Test images comprised a target person and other items digitally embedded into background scenes. Some of the background scenes and actors used in Experiments 1 and/or 2 were used again, while others were captured by the experimenter or downloaded from a licensed internet source (https://search.creativecommons.org). The downloaded photographs were marked as dedicated to the public domain (CC0) which allows copying, modifying, and distributing, all without asking permission. These images were manipulated for use as test images; this process was undertaken using the same method employed in Experiments 1 and 2 (see Section 3.2). For instance, some downloaded images needed manipulation to change sizes and/or remove people from the scene (Figure 6.1), or to crop and/or change sizes of wanted items (target persons, dogs, vehicles) to embed them into a different background (Figure 6.2).

Figure 6.1. Example of image manipulation. *(a)* Image downloaded from an online source, the person in this image was not required for the current work. *(b)* The same image after removing the unwanted person and changing the size of the image; this image was then used as a background scene.

Figure 6.2. Example of image manipulation. *(a)* Image downloaded from an online source; in this example, the person was required, but not the background. *(b)* The extracted person after being resized and embedded into a new background scene.

To generate test images, two main images were used; first, images of the target person portrayed by 20 actors (equal in gender: 10 female/male, this was for generalisability) who differed in appearance, age, clothing, and posing (Appendix L, Figure L.1). All target persons were in a pose of walking normally in the street at an interpersonal distance of 4 m. Second, images of the background scene that used 13 different images, mostly pedestrian paths with different environments such as after-dark with different exposures, and in the daytime (Appendix L, Figure L.2). The reason for using different backgrounds was that presenting the same background with different targets could make it easy to spot the target, while varying the backgrounds provides something different to look at; therefore, the task would not be obvious or predictable. Thus, the same variations were used in different backgrounds. Figure 6.3 presents an example of a test image.

Figure 6.3. Example of a test image used in the experiment which shows a target person walking in the middle of the scene.

As with Experiment 2, the images in this experiment represented a distance of 4 m between the observer and the target. This is the distance suggested first by Caminada and van Bommel (1984) based on their interpretation of Hall's (Hall, 1966, p. 123) zones of proximity, and subsequently used in other work (e. g. Rombauts et al. (1989); Johansson & Rahm, (2015)) and stated in lighting standards documents (e.g. CIE 136:2000) as the minimum distance for evaluating other pedestrians (see Sections 1.3 and 3.2).

Caminada and van Bommel assert that at 4 m evasive or defensive action is still possible: others might disagree (Fotios et al., 2016; Fotios, Yang, & Uttley, 2015). A distance of 4 m could be too close for other people, especially for the elderly, and people with impaired mobility. Subsequent research has disputed the 4 m assumption and suggests instead an interpersonal distance of 15m. This was found using eye tracking, with 15 m being the distance at which other pedestrians are first fixated (Fotios, Yang, & Uttley, 2015). At that distance there is greater time for evasive or defensive action.

An interpersonal distance of 4 m was used in Experiment 3 for convenience in the creation of the test images, since actors in Experiment 2 were modified to fit at that distance. By considering also an area-weighted analysis of the fixation distributions the question of target size is overcome to some extent. In further work it would be useful to repeat this experiment but with different interpersonal distances.

Test images were presented with all possible combinations, and randomised using bespoke software based on a Python script that was written by Dr Chris Cheal. Every variable object (e.g., target person) had its own image which was transparent except for the object (see Appendix L, Figures L.1, L.3, and L.4), and multiple images of objects could be layered on top of a background (see Appendix L, Figures L.2 and L.5). To illustrate, the area surrounding the target person (for example) was transparent, so when it was layered (pasted) on top of the background image, those (100%) transparent areas would display the background.

The software was designed to generate images of all possible combinations between test variables (target person, backgrounds, and distractions) and then randomly pick some images and layer it to present a test image. Consequently, thousands of composite images could be made (by layering different images on top each other) from all possible combinations of backgrounds, target persons, and any other objects. Importantly, the photographs were chosen to ensure the layering positions were realistic when added together. For instance, the position of target persons on the backgrounds should not show them walking into a wall, nor should the images overlap.

All images were the same size, and the targets were in the correct positions relative to the border. Overall, 18 test images were used after being layered from different categories (see Table 6.3), and presented randomly in a certain sequence, with a three-seconds observation time for each image.

6.2.1. Analysed images

The eye-tracking analysis aimed to compare fixations on primary images against other images. The primary images depicted a typical person walking normally in the middle of the street. The other images displayed the same primary images but contained distracting item(s), or a modification to the person by shifting them to the right or to the left side, or obscuring his/her face or body. The purpose of these test images was either to check the validity of the experiment or assess the research hypotheses. Therefore, it were divided into three categories: primary images, control images, and obscuring face & body; Table 6.1 details the number of images used to composite these images for each category. Next, each category of the test images is discussed in detail.

Table 6.1. The compositing of test images used in the eye-tracking – Experiment 3.

Category of test images	Subcategory	Number of images used to generate the target person	Number of images used to generate the background
Primary	-------	20 (10 female/male)	13
Controls	Left position	20 on left	13
	Right position	20 on right	13
	Dog	6 with a dog* (3 female/male)	13
	Pair of people	20 (10 female/male) + 8 pairs of people	13
	Vehicles	20 actors (10 female/male) + vehicle(s)	7**
Obscuring face & body	Obscured face	20 actors with obscured face	13
	Obscured body	20 actors with obscured body	13

* The six target persons differed from the main 20 actors because it was not efficient or realistic to embed a dog beside each of the 20 actors.
** The seven backgrounds differed from the main 13 backgrounds because it was not efficient or realistic to embed vehicles into each of the 13 backgrounds.

For the primary images, they were composited using all possible combinations between 20 actors (Appendix L, Figures L.1) and 13 backgrounds (Appendix L, Figures L.2), giving a total number of 130 combinations per gender. The experiment did not present all these combinations to every participant, as each of them observed one combination per gender chosen randomly and shown once; thus, two images were observed as primary images. Figure 6.4 presents examples of the primary images used in the experiment.

Figure 6.4. Examples of primary images used in the experiment show a target person walking in the middle of the street. *(a)* After-dark background scene with female target. *(b)* Daytime background scene with male target.

For control images, it was hypothesised that people would look at the person's face which was tested by the primary images. The challenge in the control images was to alter them to stop somebody from looking at the person. To achieve this, three variations were added/changed in the primary images to see whether the fixations would change; first, changing the target's position; second, showing only the background with no target (non-analysed images, see Section 6.2.2); and third, scattering distraction(s) in the scene such as a pair of people, animal, and vehicles. These control images were compared with the primary images to check validity and to determine whether the fixation focus differed when one of these distractions was present. For instance, if changing the target's position or adding the dog led to a different fixation, then it had an influence.

For the left position and right position images were used to test the influence of the actors' position. Assuming people looked at the person (anywhere on the person), the question that arises is, did people look at the person because they were in the middle? To offset this, the position of the target person on the footpath changed randomly to either left or right. These images were composited using all possible combinations between 40 actors (20 on the left, 20 on the right) and 13 backgrounds, giving a total number of 260 combinations per gender for each side. The experiment did not present all these combinations to every participant; each of whom observed one combination per gender for each side chosen randomly and shown once; thus, a total of four images were observed as position images. Figure 6.5 presents example of the left position and right position images used in the experiment.

Figure 6.5. Examples of position images used in the experiment showing a target person not walking in the middle of the street. *(a)* A female target on the left side. *(b)* A male target on the right side.

For the distraction (i.e. control) images, they were applied to test the influence of certain distractions in the scene such as a pair of people, dog, and vehicles. Assuming people looked at the target person, the question that arises is, did people look at the person (or face) because there was nothing else to look at? To offset this, three kinds of distractions were added separately into the primary images.

The first distraction was a pair of people (one of each gender) in a walking away position (facing backwards) behind the foreground target. These images were composited using all possible combinations between the primary images (20 actors × 13 backgrounds) and eight images of paired people (Appendix L, Figure L.3), giving a total of 1040 combinations per gender of the foreground target. The experiment did not present all these combinations to every participant, each of whom observed one combination per gender of the foreground target chosen randomly and shown once; therefore, a total of two images were observed as pair of people images. Figure 6.6 presents examples of the pair of people images used in the experiment.

Figure 6.6. Examples of pair of people images used in the experiment. *(a)* A male target with other people walking away. *(b)* A female target with other people walking away.

The second distraction was a dog, where an image of a dog walking forward beside the target person was added into the scene. These images were composited using all possible combinations between six actors (equal in gender) with a dog (Appendix L, Figure L.4) and 13 backgrounds, giving a total number of 39 combinations per gender. The experiment did not present all these combinations to every participant; each of whom observed one combination per gender, chosen randomly and shown once; thus, a total of two images were observed as dog images. Figure 6.7 displays examples of the dog images used in the experiment.

Figure 6.7. Examples of dog images used in the experiment. *(a)* A male target with a dog. *(b)* A female target with a dog.

The third distraction was vehicles, which different vehicle(s) in mixed places were added around the person target. These images were composited using all possible combinations between 20 actors and five different backgrounds with vehicle(s) (Appendix L, Figure L.5), giving a total of 50 combinations per gender. The experiment did not present all these combinations to every participant, each of whom observed one combination per gender, chosen randomly and shown once; thus, a total of two images were observed as vehicle images. Figure 6.8 presents examples of the vehicle images used in the experiment.

Figure 6.8. Examples of vehicle images used in the experiment. *(a)* A male target with vehicle. *(b)* A female target with different vehicles.

For the obscuring face & body images, they were used to check the research hypotheses. Assuming people looked at the face of the target person, the question that arises is, did people look at the face more than (or before) the body? What happened when the face or body was obscured? If the face was removed, were the fixations that would have been on the face transferred to the body or to the background?

The control images tested the question of fixation towards the face versus fixation towards the middle of the scene regardless of content by moving the actors to the left and right-hand sides of the image. If the face is the important cue, then fixations would be towards a face at the side of the image rather than to the middle of the image. A further consideration is the degree to which fixation on the face is cued by its position above the body: this was tested through the inclusion of images in which there was a face but the body was obscured (see Figure 6.9 (b)). Observations of these images would show the tendency to fixate on faces without that fixation being cued by the body. For completeness, there were further images of the body but with the face obscured.

The obscuring face & body images were the same primary images but with the face or body of the person removed. These images were composited using all possible combinations between 40 actors (20 obscured face, 20 obscured body) and 13 backgrounds, giving a total of 260 combinations per gender for each of the obscured face and obscured body images. The experiment did not present all these combinations to every participant, each of whom observed one combination per gender for each obscured face and body, shown randomly and once; thus, a total of four images were observed as obscuring face & body images. Figure 6.9 presents example of the obscured face and obscured body images used in the experiment.

The method for obscuring the face and body was derived from previous studies (Burton et al., 1999; Hahn et al., 2016). This involved using a black or a blurred square on the target face and body (see Appendix B, Table B.2). The final method of obscurement was chosen according to whether it was realistic; for instance, a black or a blurred square is not something people see on other pedestrians.

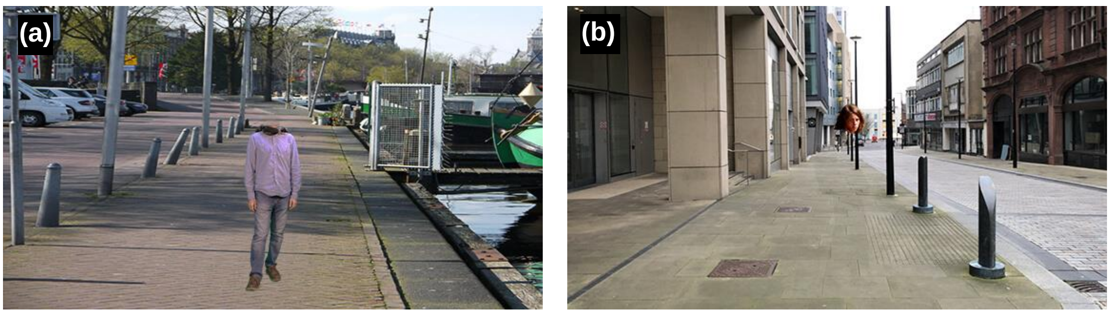


Figure 6.9. Examples of the obscuring face & body images used in the experiment. *(a)* A male target with obscured face. *(b)* A female target with obscured body.

6.2.2. Non-analysed images

There were other images used to run the experiment, where the target person in these images was not included, and they were not analysed. These images were divided to four categories: instructions images, calibrations images, non-specific images, and responses images (for Block 2 only, see below Section 6.4). Table 6.2 details the number of images used for each category. The next part will discuss separately each category of these images; except the instructions and the response images that are described in Section 6.4.

Table 6.2. The category images used in running Experiment 3, along with the number of images used for each category.

Image category	Number of images used in the image category
Instructions	2 text images
Response	2 text images
Non-specific images	20 random scenes (10 indoor/outdoor)
Calibrations	21 images with object (4 objects at all corner & centre + blank screen)

The non-specific images were random scenes shown before each test image. The reason for including them was: (1) to prevent the test images from becoming repetitive; (2) to randomise the fixations of the participant, and (3) to remove the previous image from their memory. These images were categorised into two groups: indoor scenes and outdoor scenes. Figure 6.10 presents example from each group.

Figure 6.10. Examples of non-specific images used in the experiment. *(a)* Example of indoor scene. *(b)* Example of outdoor scene.

The calibration images were used to accurately measure the location of the gaze point by determining the centre of the user's pupil, thus ensuring the fixation mark (dashed circle shown on the eye-tracking software, see Figure 6.16) was in the exact place where the participant was looking at.

These images were shown before every test image (not at the beginning of the procedure as Fletcher-Watson et al., (2008)). The reasons for doing this were as follows: (1) there could be potential movements while wearing the glasses that might affect the calibration by shaking the head, including slight movements such as scratching the head or heavy movements such as sneezing and/or knocking the glasses; (2) it might be necessary during the analysis to characterise which calibration image was shown, or to know the location of the previous fixation point that was seen before the test image. The calibration was watched by the experimenter during the process, enabling him to see where the participant was looking and make adjustments where necessary.

These images contained different objects as a target (Figure 6.11). Each object was used to generate the calibration images by allocating it in five fixed places (the four corners and the middle) on a grey background. This gave a total of 20 images (4 objects × 5 places) in addition to a blank screen used before each calibration image. Figure 6.12 presents examples of the calibration images.

Using these objects as targets beside the cross sign (the most commonly used, and the default sign on the calibration system) was a precaution to avoid two crosses (calibration object + fixation mark), where it could be hard to discriminate between them. However, the fixation mark used in the software to indicate the gaze place was a dashed circle not a cross.

Figure 6.11. The objects used as targets in the calibration images. *From the left:* eye, clock, apple, and cross.

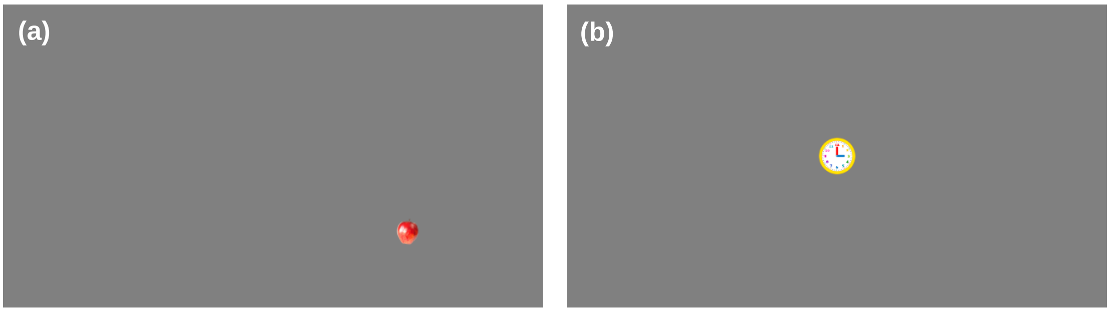


Figure 6.12. Examples of calibration images used in the experiment. *(a)* An apple displayed at the bottom right of the image. (*b)* A clock displayed in the middle of the image.

Additionally, two other categories were generated for further exploration of the effect of the background, but these were not analysed due to time constraints. These were images of only-background of a scene with no targets, and images of only green background, the following will discuss each one.

The only-background images were used to test the influence of the scenes. Assuming people looked at the target person, the question that arises is, did people look at the person (or face) because that one target was interesting? To offset this, the target was removed from the scene. These images used the same 13 background images with no embedded target person or distraction items in the scene. The experiment did not present all the 13 backgrounds to every participant, each of whom observed two backgrounds chosen randomly and shown once; thus, a total of two images were observed as only-background images. Figure 6.13 presents examples of only-background images used in the experiment.

Figure 6.13. Examples of only-background images used in the experiment, showing only a scene without any target. *(a)* A daytime scene. *(b)* An after-dark scene.

For the green-background image, it was used to check the background effect in this experiment by using this uniform neutral background. This image was used as a benchmark to compare fixations with real background scenes. Thus, in each category from the test images, the green-background was used once. For instance, if two primary images were observed, this means one of them used the green-background. Figure 6.14 presents some examples of test images with the green-background.

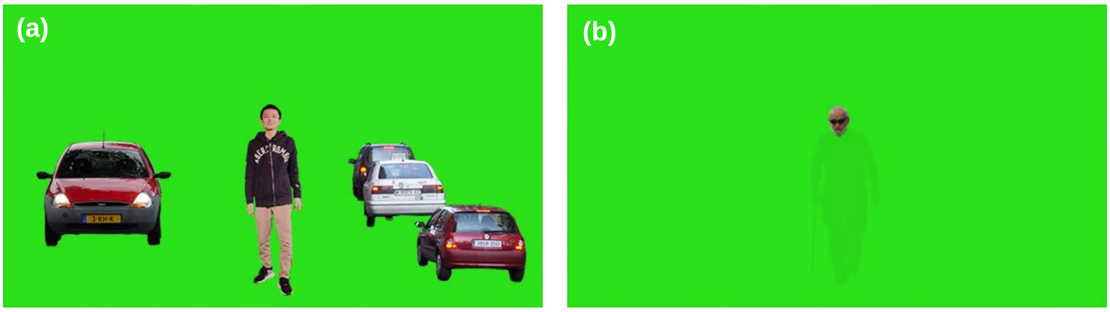


Figure 6.14. Examples of test images with the green-background used in the experiment. *(a)* A control image shows a male target with vehicles. *(b)* An obscuring face & body image shows a male target with an obscured body.

6.3. Eye-tracking apparatus

A series of test images were observed by test participants while wearing an eye-tracking device (SMI ETG 2W analysis Pro). This device is a pair of spectacles (worn as a normal pair of glasses) with inbuilt cameras that record the user's eye movement in the viewed scene (Figure 6.15). The output data is a video of the viewed scene with the fixation mark (dashed circle) on the fixation place that indicates where the user was looking at; the video was recorded onto the laptop using bespoke software (iViewETG).

Figure 6.15. The SMI eye-tracking glasses used in the experiment.

The collected data (recorded video) were then analysed using BeGaze 3.7 software. By following the manufacturer's instructions for using the eye-tracking device; the analysis utilised a semantic mapping feature in order to make fixations to be meaningful. To illustrate, the analysis maps each fixation with a reference describing what (and where) the participant was looking at. Thus, for this semantic mapping stage, eight specified domains were used: face, body, background, dog, other people, vehicles, obscured face, and obscured body.

Table 6.3 details the categories of the test images with the analysed domains for each category. As an example of this processing, when the fixation is on the body, the experimenter classified it as being on the body (Figure 6.16). After classifying all fixations, Begaze extracted these data to Excel for further analysis (see results in Chapter 7).

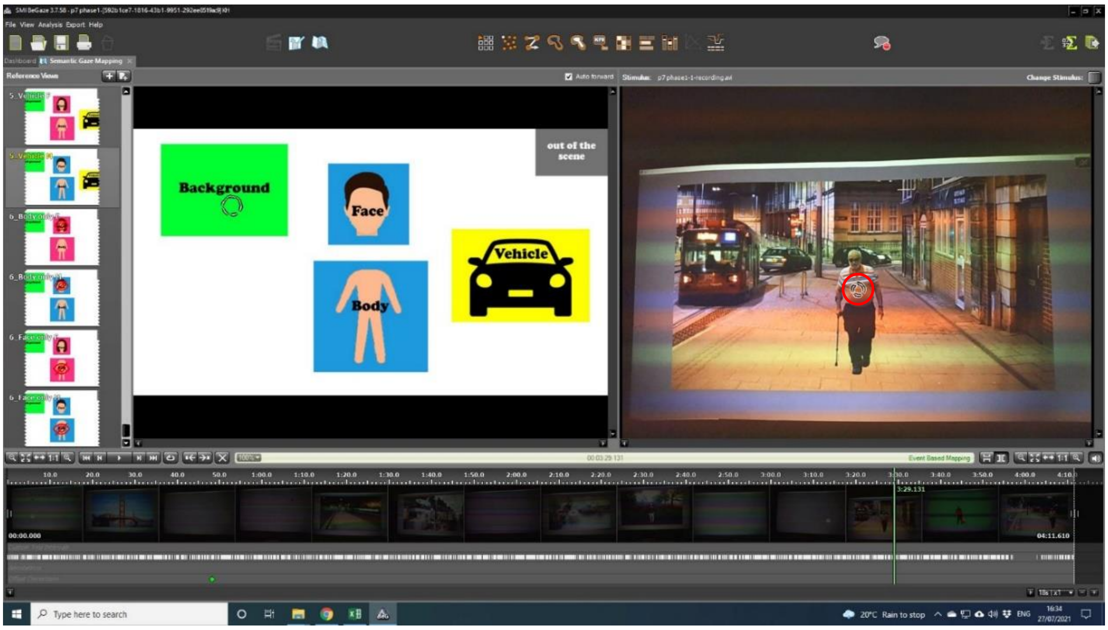


Figure 6.16. Screenshot of the semantic mapping analysis from the Begaze software which shows two windows. *(Right)* the scene viewed by the participant during the trial with the fixation mark (dashed circle; inside the red circle); here, the fixation is on the body of the target. *(Left)* the assigned domains created by the experimenter in order to link each fixation to a corresponding domain.

Table 6.3. Description for each category of the 18 test images with the analysed domains for each category.

<table>
<tr><th>Category of test images</th><th>Subcategory</th><th>n</th><th>Purpose</th><th>Description of the image</th><th colspan="4">Test domains on the image:</th></tr>
<tr><td>Primary</td><td>-------</td><td>2</td><td>Compare it with other images</td><td>One person walking in the middle of the street</td><td rowspan="6">Face</td><td rowspan="6">Body</td><td rowspan="6">Background</td><td>-------</td></tr>
<tr><td rowspan="5">Control</td><td>Left position</td><td>2</td><td rowspan="5">Check the validity of the experiment</td><td>One person walking on the left side of the street</td><td>-------</td></tr>
<tr><td>Right position</td><td>2</td><td>One person walking on the right side of the street</td><td>-------</td></tr>
<tr><td>Pair of people</td><td>4</td><td>One person walking in the middle of the street with two other people between him/her</td><td>Other people</td></tr>
<tr><td>Dog</td><td>2</td><td>One person walking in the middle of the street with a dog</td><td>Dog</td></tr>
<tr><td>Vehicles</td><td>2</td><td>One person walking in the middle of the street with vehicle(s) in the scene</td><td>Vehicle(s)</td></tr>
<tr><td rowspan="2">Obscuring face & body</td><td>Obscured face</td><td>2</td><td rowspan="2">Check the research hypotheses</td><td>One person with obscured face walking in the middle of the street</td><td>Obscured face</td><td>Body</td><td rowspan="2">Background</td><td>-------</td></tr>
<tr><td>Obscured body</td><td>2</td><td>One person with obscured body walking in the middle of the street</td><td>Face</td><td>Obscured body</td><td>-------</td></tr>
</table>

The test domains were varied in size, among which the face presented the smallest area. Thus, if the fixation locations had been randomly distributed, this would have favoured large areas such as the background or body. Therefore, an area-weighted analysis was conducted where the fixation durations in each domain were weighted according to the percentage of the whole image. The domain areas were measured using the "area" command in AutoCAD software (Figure 6.17). Consequently, the duration of fixations on each domain was divided by the percentage of each domain across all images. For example, fixations on faces were divided by 0.42%, and on bodies by 4.67% (Table 6.4).

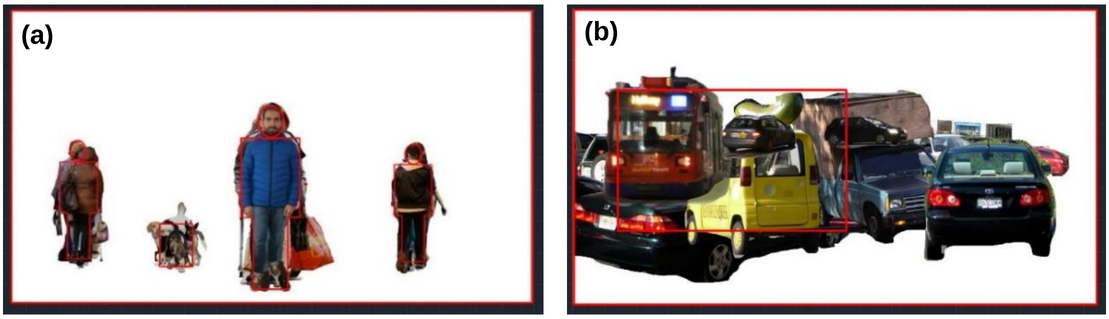


Figure 6.17. Measuring test domains using AutoCAD, where the red lines (rectangle and circle) represent the area of the domain across all images. *(a)* Estimating the size of the background, face, body, dog, and pair of people. *(b)* Estimating the size of the vehicles, where the red rectangle denotes the average size of all vehicles.

Table 6.4. The percentage of each domain used in the area-weighted analysis; this was measured using AutoCAD software.

Category of test images	Subcategory	Test domains on the image:					
		Face	Body	BG****	Pair of people	Dog	Vehicles
Primary	-------	0.42%	4.67%	94.91%	---	---	---
Control	Left position	0.42%	4.67%	94.91%	---	---	---
	Right position	0.42%	4.67%	94.91%	---	---	---
	Dog	0.42%	4.67%	93.95%	---	0.97%	---
	Pair of people*	0.42%	4.67%	90.68%	4.23%	---	---
	Vehicles	0.42%	4.67%	74.16%	---	---	20.76%
Obscuring face & body	Obscured face**	0.42%	4.67%	94.91%	---	---	---
	Obscured body***	0.42%	4.67%	94.91%	---	---	---

* Across left and right persons.
** Face was transparent, included only the area where the face should be.
*** Body was transparent, included only the area where the body should be.
**** Background

6.4. Procedure

In the laboratory, the light setting was the same as that used in Experiment 2 (see Section 3.4). In this experiment, the vertical illuminance at the observer's eye was 23 lx (mean across all test images), and the average horizontal illuminance on the desk in front of the participant was 25 lx; these measuring points are shown in Figure 6.18. At the start of the session, an adaptation time of 20 minutes was given to adapt to the low light level together with a brief explanation of the test. While this time, participants were asked to sign the consent form and read the participant information sheet (if required). Examples of a participant information sheet and consent form can be found in Appendix E, Figures E.5 and E.6. The test participants sat on a chair while wearing eye-tracking glasses, with a glasses strap to hold the spectacles in place. Participants were seated two metres away from the screen, and observed the images whilst wearing the eye-tracking glasses (Figure 6.18).

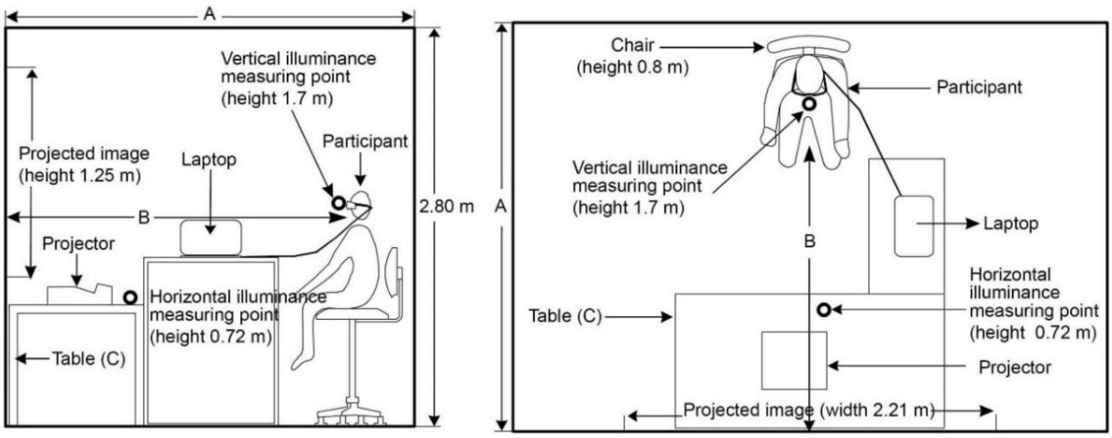


Figure 6.18. Diagram of the apparatus in Experiment 3: section *(Left)* and plan *(Right)*. Dimensions: A = 2.43 m (laboratory length from front to back wall). B = 2 m (from the projected image to the participant). C = table dimensions; L = 1.6 m, W = 0.80 m, H = 0.72 m. The vertical centre of the projected image was aligned with the typical eye level of a standing adult.

The experiment included two blocks of trials differentiated by the task set. In Block 1 (free viewing condition), the task was simply to observe the images; in Block 2 (safety rating determination), participants were required to respond to a question after each image. In both blocks, the images (i.e. test domains, e.g. target person, background, and vehicles) were different for each participant. The order of the two blocks was randomised with balance (alternate method). For example, half of the participants completed Block 1 first, and the second half completed Block 2 first, but in an alternate manner; for example, if Participant 1 completed Block 1 first, Participant 2 has to complete Block 2 first, then Participant 3 has to complete Block 1 first…and so on. This way of randomisation made it possible to analyse the differences by comparing the results in Block 1 when it was in first (and second) place against the results of the same block when it was in the other place; however, an analysis was conducted on the two blocks (whatever the order) to see their effect regardless their order. This randomisation was performed manually by the experimenter using a checklist that was created to: (1) check the order of randomisation; (2) to ensure each participant fulfilled all the requirements such as vision tests, and (3) to record notes. This checklist can be found in Appendix D, Figure D.2. The next part will describe separately each block.

Block 1 was a free viewing condition where test images were observed with no other specific task to do, just simply looking at the images. The procedure started by displaying first the instructions image where it was shown to the participant and read aloud by the experimenter (Figure 6.19).

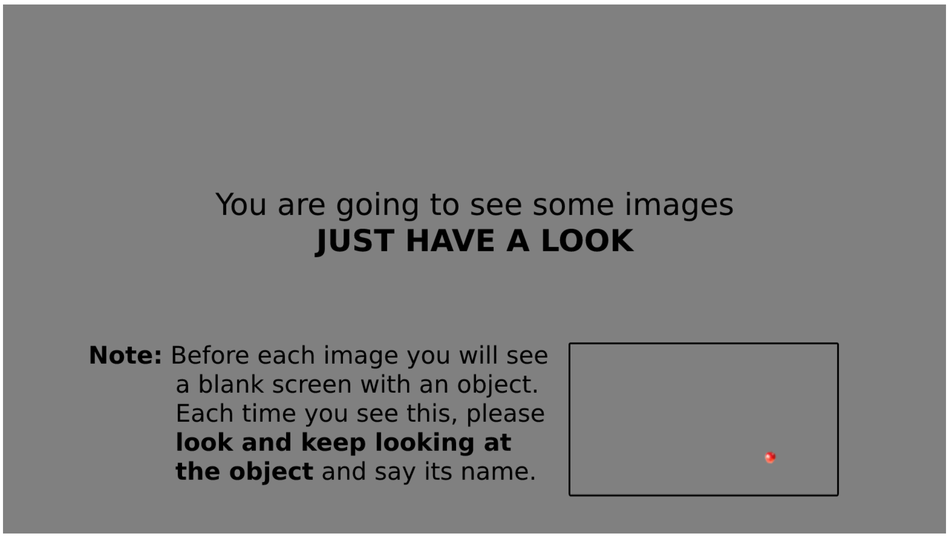


Figure 6.19. The instructions image of Block 1 (free viewing condition), showing the instructions that was presented to the participants.

Following the display of the instructions image, the process of calibrating the eye-tracking occurred by presenting an image that displayed the four objects that used in the calibration images, and then asked the participant to look at each object and say its name (Figure 6.20). This helped to introduce the object to the participant (in case they were not familiar with it) and to double check that the eye-tracking glasses were calibrated correctly. Furthermore, if a participant said the name of the object but the fixation mark was not on that object (the experimenter was able to see where their fixation mark was), then something was wrong that required the experimenter to fix it.

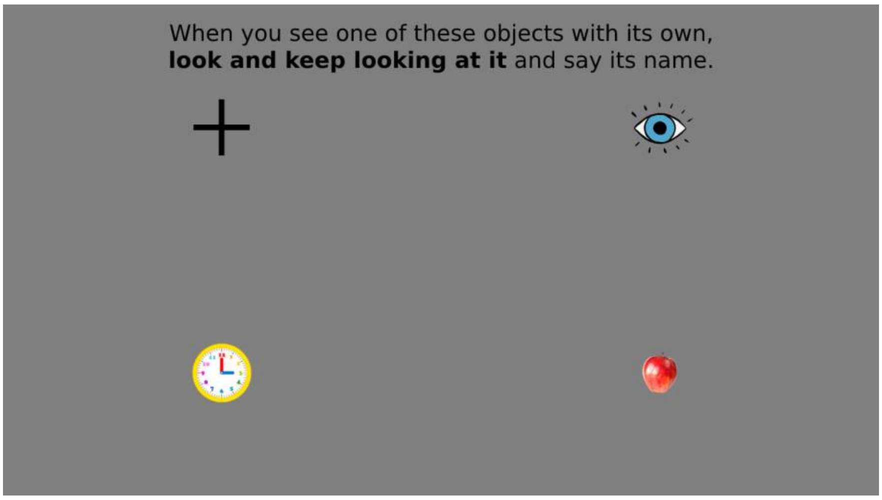


Figure 6.20. The calibration instructions image used in Blocks 1 and 2.

Block 1 presented 36 images (18 test images + 18 control images), each of which was presented for three-seconds in a random order. Before each image, there were calibration images, which were a one-second of a blank screen followed by unlimited time of a screen containing an object. For calibration, the participant was required to say the name of the object, and the experimenter confirmed the calibration is achieved by pressing a space bar on the computer to display next image. Figure 6.21 shows an example of the sequence of the images being presented in this block. There were two practice images at the beginning of the trials. This block took approximately four minutes to complete.

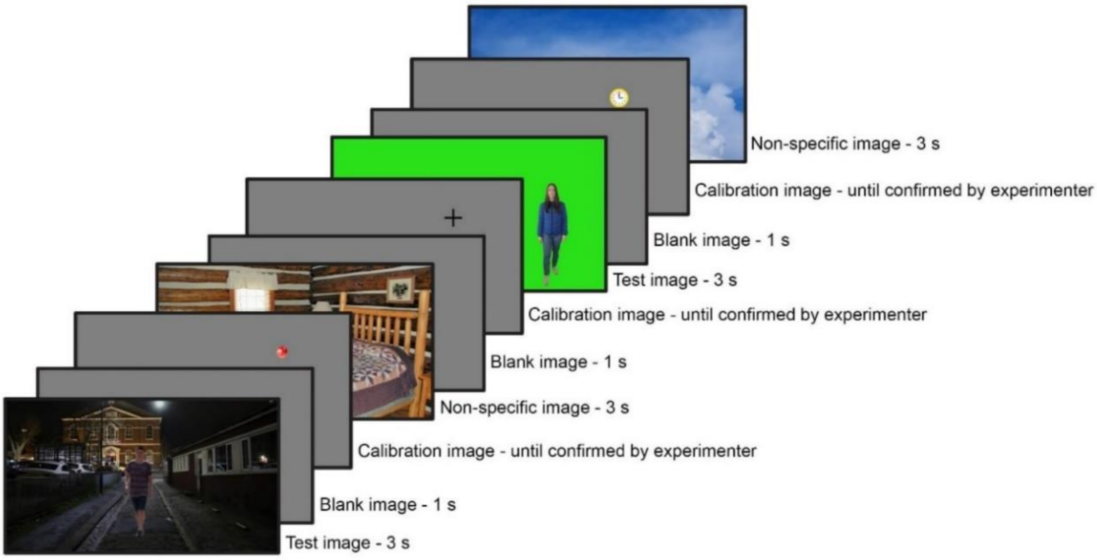


Figure 6.21. Examples of the sequence of the images in Block 1. Participants observed test images (primary, control, and obscuring face & body) in a random order; before each test image, there were calibration images, and after that a non-specific image (indoor or outdoor).

Block 2 was designed to provide a focus (question) to see whether this affected visual fixations. Test images were observed along with a question to think about while viewing the images, followed by a simple response after each image. The responses to the questions were not analysed. The procedure started by displaying first the instructions image where it was shown to the participant and was read aloud by the experimenter (Figure 6.22).

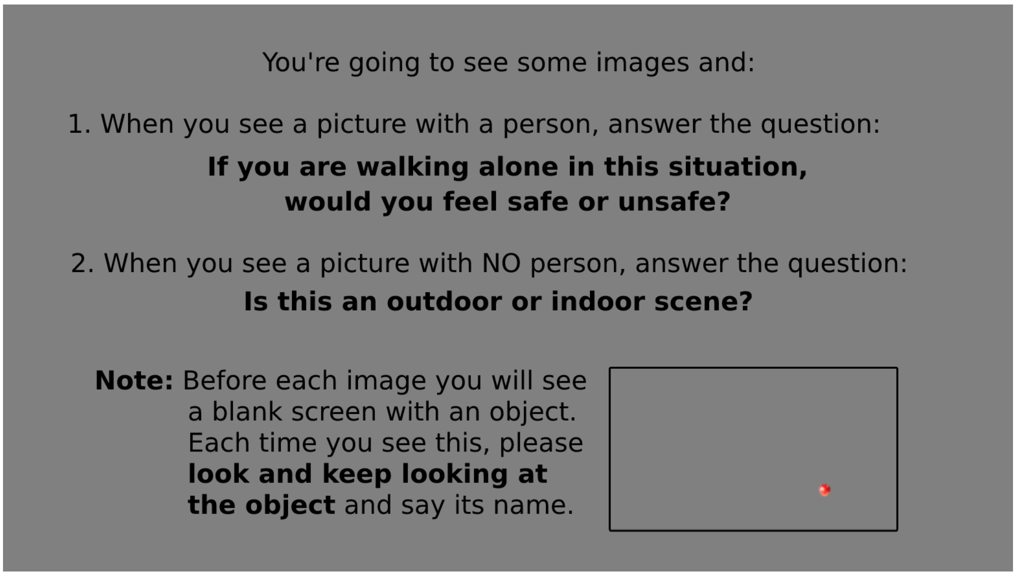


Figure 6.22. The instructions image of Block 2 (safety rating determination), showing the instructions that was presented to the participants.

Following the display of the instructions image, the process of calibration was conducted (same as Block 1), then presentation of 36 images (18 test images + 18 control images), two practice trials, and two response screens (Figure 6.23). Figure 6.24 shows an example of the sequence of these images being presented. The response images were presented after each test and displayed until the participant made a response. Participants responded orally by saying "safe” or “unsafe" and “indoor" or "outdoor”, and the experimenter then recorded that response. The reason for this categorisation was to stop the participant to anticipating the testing factors (e.g. face and body), as this would have encouraged them to look at the person (specifically the face). This block took approximately six minutes to complete. The whole experiment including the two blocks took approximately 30 minutes to conduct for each participant.

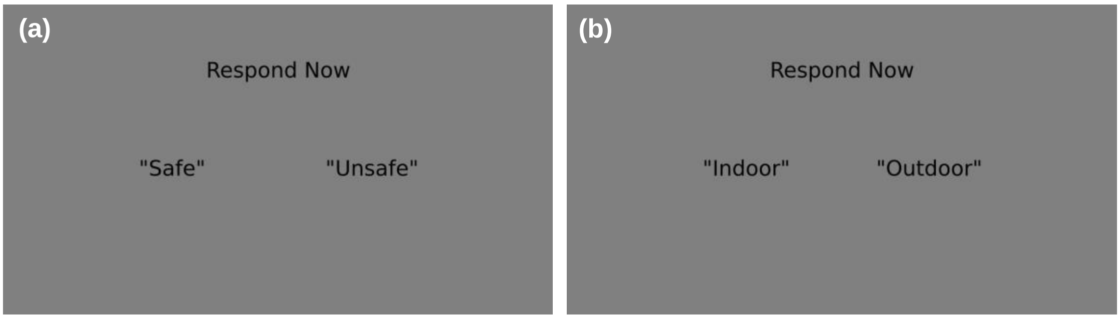


Figure 6.23. The response images used in Block 2 to answer the questions shown in Figure 6.22. *(a)* Responses for question one. *(b)* Responses for question two.

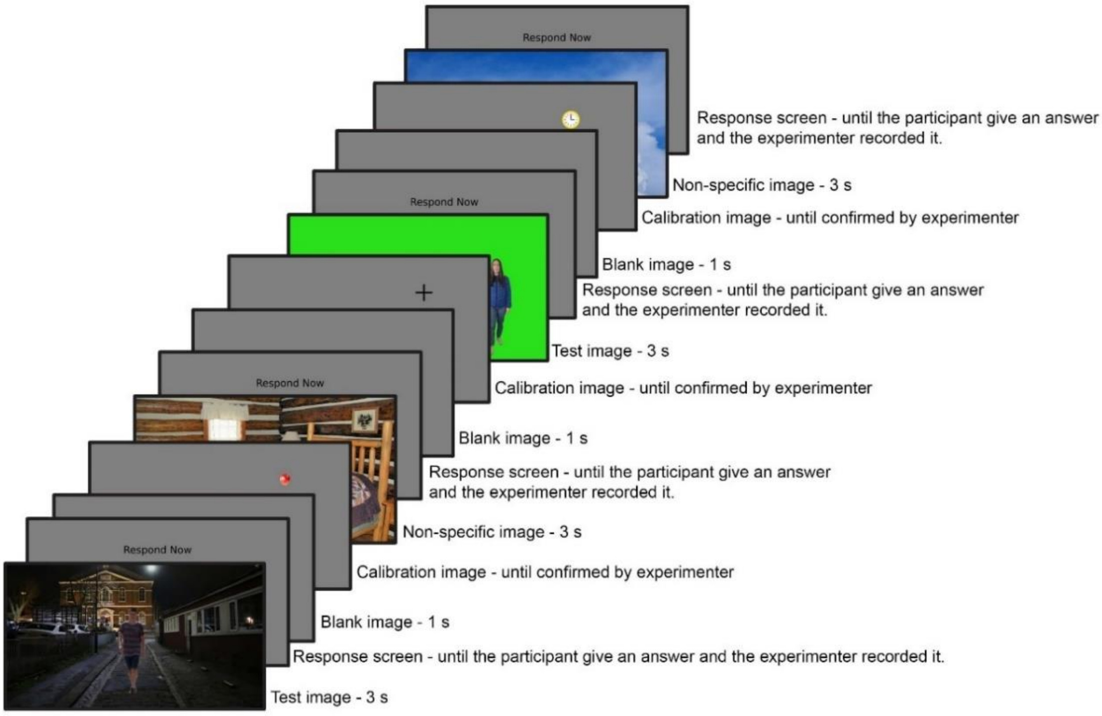


Figure 6.24. Examples the sequence of the images in Block 2. Participants observed test images (primary, control, and obscuring face & body) in random order followed by the response screen; before each test image, there were calibration images, and after that a non-specific image (indoor or outdoor) followed by the response screen.

6.5. Participants

This experiment was granted ethical approval by the University of Sheffield (reference: 037846 / 05 May 2021). A total of 32 participants were then recruited, with an equal balance of males and females. This number of participants was chosen according to Cohen's rule of thumb for a sufficient effect size (Cohen, 1992). Participants ranged between 19 and 42 years old, with most being aged between 18 and 24 (63%). They provided written consent in accordance with the protocols of the ethical approval. Those considered vulnerable were not targeted in the research. Each participant received an incentive payment of £10 for their contribution. This was in recognition of the time and commitment provided by the participant, and to cover any travel or other expenses that may have been incurred in order to attend the experimental session.

It was important to address questions of generalisability to the greater population, and to determine whether age and gender matter. Therefore, participants' age and gender were recorded to assess how well the collected data represented the whole population.

Participants were given the consent form and the information sheet via email, which outlined all the details of the experiment before they came to the laboratory. Once they agreed to participate, they received details of further requirements and how to access the laboratory. An example of a participant information sheet and consent form are provided in Appendix E, Figures E.5 and E.6.

Similar to the safety evaluation experiments, each participant conducted the experiment on an individual basis. They were required to have reasonable eyesight, including normal colour vision and normal or corrected-to-normal visual acuity, as screened at the beginning of the experiment (see Section 3.5). No participants were rejected on the basis of inadequate visual acuity or colour vision.

Because the experiment required the wearing of eye-tracking glasses, it was not possible to recruit people wearing glasses at the time of the experiment. This was specified when recruiting participants and mentioned in the information sheet. There was no problem if they wore contact lenses, or were happy to take off their spectacles and could still see adequately.

This experiment was conducted in June 2021 during the coronavirus pandemic (COVID-19). Consequently, an additional document was sent within the participant information sheet and consent form; this was a separate document detailing a preventative plan with regard to the pandemic and appropriate steps to prevent contagion (see Appendix E, Figure E.6).

6.6. Summary

This chapter described the eye-tracking method used in Experiment 3 to investigate visual cues and to measure visual fixations towards other people (i.e. face and body) and other items (e.g. background and distractions such as dog). A previous study by Fletcher-Watson et al. (2008) investigated gaze behaviour, but did not include a validation procedure. Experiment 3 repeated their experiment but overcame some of the limitations such as building validation steps and using full-size target persons rather than cropped person which might hide certain features on the body.

Experiment 3 presented a series of test images to measure the location of first fixations and the duration of fixations towards the face, body, background, and distraction items (dogs, other people, and vehicles). Distraction items were used as a validation step to ensure participants did not look at the person immediately. The images were observed under two test blocks: Block 1 involved free viewing without any other required task (just have a look); Block 2 required participants to answer a question about the observed image in order to see whether fixation changes when people have safety thoughts in mind.

The next chapter presents the results of the eye-tracking – Experiment 3, reporting the duration of fixations towards specified domains and the location of first fixations. These results are then compared with the results of Experiments 1 and 2 to form the conclusion of this thesis.

Chapter 7. Eye-Tracking – Experiment 3: Results

7.1. Introduction

Chapter 6 described the use of eye-tracking method of Experiment 3 conducted to assess the nature of visual fixations on other people. Eye-tracking provides an objective measure of visual attention, and hence an alternative to the subjective evaluations used in Experiments 1 and 2. This chapter presents the results and analysis of Experiment 3 to answer the question: is the face an important visual cue?

Critical fixations are assumed to be important for pedestrians, and a priority feature for road lighting after-dark. This experiment extended past work on eye movement (Crouzet et al., 2010; Fletcher-Watson et al., 2008; Guyonneau et al., 2006; Kirchner & Thorpe, 2006) by repeating the experiment conducted by Fletcher-Watson et al. (2008), but with the inclusion of additional validation steps (see Section 6.1). The results were analysed following the same approach adopted by Fletcher-Watson et al. (2008) which was to measure the duration of all fixations and the location of first fixations on different domains of the scene (e.g. face, vehicle). This was done by analysing the recorded video (using a semantic mapping feature: see Section 6.3) of each test image containing a person across all participants, and then calculating the percentage time spent on different domains of the scene, and on the first fixations.

7.2. Analysis of fixations

The data for this experiment were collected using eye-tracking glasses which recorded the eye movements of participants and were extracted as a recorded video of the viewed scene, with a participant's gaze point displayed as a dashed circle (see Figure 6.16). All visual fixations were extracted and analysed using eye-tracking software (BeGaze 3.7). The analysis specified each fixation towards one of the test domains, as described in Section 6.3.

Although 32 people participated in the experiment, an error in data recording led to the removal of responses from three participants. This was due to an unknown technical issue (the SMI eye-tracking equipment and BeGaze manual did not provide reasons for this) where the responses were recorded in a low tracking ratio that caused incorrect location of the fixation mark and an overlap in recording the video and audio. Similar issues have been reported in previous eye-tracking studies in which the samples were also reduced (Mantuano et al., 2017; Vansteenkiste et al., 2014, 2017). The generation of incorrect data, or losing partial or full data, is an acceptable reason for excluding trials from analysis (Holmqvist et al., 2011).

Data distribution was tested using the method described in Section 4.2. Overall, the data distribution analysis of all fixations assigned to the domains suggested they were not drawn from normally distributed populations (see Appendix H, Table H.5).

Each participant completed two blocks of trials: Block 1 involving free viewing with no other task; Block 2 required a respond to a question after each image (see Section 6.4). In each block, the test participants observed the 18 test images, each displayed for three-seconds.

The 18 test images included three categories labelled ‘primary’, ‘control’, and ‘obscuring face & body’ (see Table 6.3). These 18 images were: two ‘primary’ images showing a single person against a background; 12 ‘control’ images (to verify conclusions drawn from the ‘primary’ images) showing a single person against a background with one of the following distractions: change in the position of the person (to the left or right hand side of the image, rather than being central), and the presence of other items: 'pair of people', 'dog', and 'vehicles'; and four ‘obscuring face & body’ images of a single person with their face or body obscured. The aim was to compare fixations on each of the three main domains: face, body, and background; as well as three other domains (as distractor): people, dog, and vehicles.

One problem when interpreting gaze direction from eye-tracking videos is when the fixation mark does not clearly fall into a specific domain. Figure 7.1 shows an example image where gaze direction was uncertain: here the researcher could not be reasonably confident of recording the correct domain and it was hard to make a decision. In the current work, the researcher noted nine such images. In these cases, the researcher imagined a dot in the centre to make the decision. For example, the decision for Figure 7.1 (a) was background, and Figure 7.1 (b) was body.

Figure 7.1. Examples of images where the fixation mark (dashed circle) not clearly identify the domain into which gaze falls. *(a)* The fixation mark could indicate gaze towards the body, the vehicle, or the background. *(b)* The fixation mark does not clearly indicate whether gaze is towards the face or body.

To validate the consistency of the experimenter's allocation of fixations to different domains, a validation exercise was conducted. This done by asking six other researchers (as validators) to review the fixations recorded in 27 images. This set of images included three groups: (1) the nine images where the experimenter was unsure about the fixation (Appendix M, Figure M.1), (2) nine where the experimenter was confident about the fixation domain, (Appendix M, Figure M.2), and (3) nine images selected randomly (Appendix M, Figure M.3). The 27 images were checked in a random order: an example of the used form is shown in Appendix M, Figure M.4. The results of this validation exercise are shown in Table 7.1. For those nine images where the experimenter was sure of fixation, and for the nine randomly chosen images, the validators agree with the experimenter's decision in 100% of cases. In those nine images where the experimenter was not confident of fixation domain, validators agree in 80% of cases.

Table 7.1. The three groups of images used in the validation exercise along with the percentage of specifying the image domains compared with what the experimenter decisions.

Image group	**Image domains recorded by the validators**	
	The same as the experimenter	**Different to the experimenter**
(1) Experimenter not confident of fixation domain	80%	20%
(2) Experimenter confident of fixation domain	100%	0%
(3) Randomly chosen images	100%	0%

7.3. Analysis of test images

7.3.1. Duration of fixations

In this analysis, the percentage distribution of fixation durations across male and female targets is presented for both the normal-size and the area-weighted scores in Blocks 1 and 2. The area-weighted analysis was assumed to represent an equal percentage of the scene for each domain, see Section 6.3. Next part will discuss these results for each block, and the significant differences between the test domains on each of the images.

Table 7.2 (a) presents the percentage of fixation durations of the normal-size for each test image in Blocks 1 and 2. In Block 1 participants were required to do a free viewing with no other task. In the results of normal-size durations, the highest percentage of fixation durations in the primary images was towards the face (41%), followed by the background (33%), and the body (26%). This indicates that when there was a target person in the middle of the scene, the face received the most visual attention.

In the control images, where the target person was not in the middle of the image but either on the left or the right side, the highest percentage of fixation durations was towards the background (left: 56%, right: 63%) with reduced attention to the target person, followed by the body (26%) if the target was on the left, or the face (21%) if it was on the right. This indicates that when the target person was not in the middle of the scene, the fixations towards them reduced. In subsequent control images where the target person was in the middle of the scene but a distraction item(s) was included. When the distraction was a dog or pair of people, the highest percentage of fixation durations was on the background (30% and 31%, respectively) rather than the face (25% and 29%, respectively), and finally the dog (18%) or the body (16%), respectively. When the distraction was a vehicle(s), the highest percentage of fixation durations was on the face (33%), followed by the vehicles (27%), and finally the background (17%). Among the distraction items used, the vehicles were the most distracting; when these were included, fixations towards the person (especially the face) were higher than when the distractions were a dog or a pair of people.

In obscuring face & body images, when the face was obscured, the highest percentage of fixation durations was on the body (42%), followed by the background (34%), and finally the obscured face (23%). By contrast, when the body was obscured, fixation durations were mostly on the face (43%), followed by the background (32%), and finally the obscured body (24%). Across all test images, the greatest duration of fixations was towards the background (33%), followed by the face (25%), and then an equal duration for the distraction items and the body (21%).

The fixation durations were then weighted by area (Table 7.2 (b)). In Block 1, for all domains, the highest percentage of fixation durations was on the face (69% – 95%), followed by the body (11% – 4%), the distraction (if any) (dog: 25%, people: 7%, vehicles: 1%), and finally the background (0.2% – 1%). The control item that most affected the duration of fixations towards faces was the dog, where fixations on faces reduced to the lowest duration (69%), and the percentage of fixations on the dog (25%) was the highest of the distraction items (i.e. people and vehicles).

In Block 2, participants were required to respond to a question after observing each image. The results were the same as for Block 1, except for two cases. First, when the distraction item was a dog, the greatest difference lay on the face (11%) where the percentage changed from 25% to 36%, and there was an equal difference for the body and dog (6%) which changed from 24% to 18% (body), and from 21% to 15% (dog), whereas the lowest difference was for the background (2%) which changed from 30% to 32%. Second, for the pair of people, the greatest difference lay in fixation durations on the background (10%) where the percentage changed from 21% to 31%, followed by the pair of people (8%) which changed from 24% to 32%, and finally the face (3%) which changed from 29% to 32%, while no change occurred for fixations on the body.

Across all test images, the greatest difference lay on the face (6%) where the percentage changed from 25% to 31%, and there was reduction in the percentage of fixation durations on the body (4%) which changed from 21% to 17%, and on the background (2%) which changed from 33% to 31%, while no changes occurred for the distractions. In other words, when people had a safety question in their head, visual attention towards faces increased, in contrast with the body and background where it slightly decreased, while there was no change for the distraction; however, these percentage changes were slight (6%, 4%, 2%). This indicates that having a safety thought on their mind slightly affected people's gaze behaviour, and that attention towards faces became higher than on bodies and other objects in scenes including dogs, other people walking away, and vehicles.

Table 7.2. Fixation durations (as percentages) across male and female targets for each domain of test image in Blocks 1 and 2. *(a)* The scores of the normal-size. *(b)* The scores of the area-weighted.

Block	Category of test images	Subcategory	(a) Normal-size*				(b) Area-weighted			
			Test domains on the image:				Test domains on the image:			
			Face	**Body**	**Background**	**Distraction**	**Face**	**Body**	**Background**	**Distraction**
Block 1	**Primary**	-------	41%	26%	33%	---	94%	5%	0.3%	---
	Control	Left position	19%	26%	56%	---	88%	11%	1%	---
		Right position	21%	16%	63%	---	92%	6%	1%	---
		Dog	25%	24%	30%	21%	69%	6%	0.4%	25%
		Pair of people	29%	16%	31%	24%	88%	4%	0.4%	7%
		Vehicles	38%	18%	17%	27%	94%	4%	0.2%	1%
	Obscuring face & body	Obscured face	23%	42%	34%	---	86%	14%	1%	---
		Obscured body	43%	24%	32%	---	95%	5%	0.3%	---
	All images	-------	25%	21%	33%	21%	82%	6%	0.5%	11%
Block 2	**Primary**	-------	45%	22%	34%	---	95%	4%	0.3%	---
	Control	Left position	27%	25%	47%	---	92%	8%	1%	---
		Right position	33%	13%	54%	---	96%	3%	1%	---
		Dog	36%	18%	32%	15%	81%	4%	0.3%	15%
		Pair of people	32%	16%	21%	32%	87%	4%	0.3%	9%
		Vehicles	46%	15%	15%	25%	96%	3%	0.2%	1%
	Obscuring face & body	Obscured face	26%	42%	32%	---	87%	13%	0.5%	---
		Obscured body	51%	13%	35%	---	97%	2%	0.3%	---
	All images	-------	31%	17%	31%	21%	88%	4%	0.4%	7%

* Percentages for the domains do not total 100% because a percentage of fixations was out of the scene and/or was not accurately measured in BeGaze, as some remained in the domain even after the test image was removed. This could be due to the small tracking ratio or a technical issue.

A t-test was performed to ascertain whether there were any significant differences between the domains. The significance level was set at $p < 0.05$ (Field, 2009, p. 331). A series of t-tests was performed on all possible pairs of domains: face vs body, face vs background, face vs dog, and dog vs body. However, running several simultaneous t-tests can cause a problem, as the probability of a significant result increases with each test run; therefore, Bonferroni *post hoc* tests were applied to correct the *p*-values. For example, when three groups of tests simultaneous were run at $p = 0.05$, the correction would be a new threshold level at $p = 0.0167$; this was calculated by dividing the original *p*-value (0.005) by the number of the test groups. Table 7.3 presents the results of these new p-values for each category of test images.

Table 7.3. Results of the Bonferroni corrections for the new *p*-values.

Category of test images	Subcategory	Test groups	Number of test groups	New level of *p*-value
Primary	**-------**	Face, body, background	3	0.0167
Control	Left position			
	Right position			
	Dog	Face, body, background, dog	6	0.0083
	Pair of people	Face, body, background, people		
	Vehicles	Face, body, background, vehicles		
Obscuring face & body	Obscured face*	Face, body, background	3	0.0167
	Obscured body*	Face, body, background		

* Face and body were obscured by being transparent; the measurement of these areas was applied on the area of where the face and body should be.

Table 7.4 presents the significant differences between the domains for each test image using the t-tests and the Bonferroni corrections applied to the normal-size and the area-weighted in Blocks 1 and 2. Next part will discuss these results for each block.

In Block 1, using the t-tests, in the normal-size and the area-weighted of the primary images, the differences between all domains were significant. In the control images of the left and right positions, the differences between background and face/body were significant, but not between face and body.

In control images where a distraction item(s) was included, for the images of the dog and the pair of people, there were no significant differences between all domains in normal-size, except between the body and background for the pair of people; in area-weighted, the differences between all domains were significant, except between the body and pair of people.

For the vehicles images, two comparisons (face vs body/background) were significant in normal-size, whereas four comparisons (vehicles vs face/body/background and body vs background) were not significant; while in area-weighted, the differences between all domains were significant. For the obscuring face & body, there were no significant differences between domains for normal-size, except between the face and body when both were obscured; in area-weighted, the differences between all domains were significant.

Using the Bonferroni correction (shaded in Table 7.4), in Block 1 of the normal-size, the differences in six out of 27 comparisons were significant (none of these differences was significant: background vs face/body in left/right positions; face vs background in vehicles; and face vs body in body only). In area-weighted, the differences between all domains were significant, except for two comparisons (dog vs face, and pair of people vs body).

In Block 2 of the normal-size, using the t-test for primary images revealed one significant difference between face and body, but not for the rest (background vs face/body). In the control images for the left and the right positions, the differences between all domains were significant, except for between the face and body on the left position.

In the control images where there was a distraction(s), for the dog, three comparisons (face vs body/dog, and dog vs background) were significant, and three (background and face/body, and dog vs body) were not. For the pair of people, three comparisons (body vs face/pair of people, and pair of people vs background) were significant, and three (background and face/body, and pair of people vs face) were not.

For the distraction of vehicles, four comparisons (body vs background, and vehicles vs face/body/background) were significant, and two (face vs body/background) were not. In obscuring face & body images, the differences between all domains were significant, except for between background vs face/body (in obscured face).

Overall, in normal-size images using the t-test (Table 7.4 (a)), nine of the 33 comparisons were significant, and 24 were not; using the Bonferroni corrections, 13 were significant and 20 were not. In area-weighted (Table 7.4 (b)), the t-test and Bonferroni corrections revealed significant differences between all domains in each test image. In general, the number of significant cases was higher in Block 2 and higher in the area-weighted images, especially when using the Bonferroni corrections.

This experiment analysed fixations on people regardless of their age, gender, or other characteristics. It did not aim to explore differences between people (particularly gender). Male and female actors were used as targets in the test images to promote generalisation and increase the sample and diversity of test images.

Nevertheless, Appendix N, Tables N.1 to N.4 presented the results for each gender in Blocks 1 and 2 with respect to fixations (percentage associated with mean duration and standard deviation). In addition, Appendix N, Tables N.5 and N.6 presented the significant differences between test domains (t-test and Bonferroni) for each gender in normal-size and area-weighted in Blocks 1 and 2 respectively.

Table 7.4. Significant differences between the test domains for each test image in Blocks 1 and 2; using t-tests (p = 0.05) and Bonferroni corrections (*p* = 0.02, *p* = 0.009) (shaded), which were applied to the scores of the normal-size *(a)* and area-weighted *(b).*

Block	Category of test images	Subcategory	(a) Normal-size						(b) Area-weighted					
			Comparisons between the test domains						Comparisons between the test domains					
			Face v Body	Face v Bg.*	Body v Bg.*	Dis.** v Face	Dis.** v Body	Dis.** v Bg.*	Face v Body	Face v Bg.*	Body v Bg.*	Dis.** v Face	Dis.** v Body	Dis.** v Bg.*
Block 1	**Primary**	-------	<0.001	<0.001	<0.001	---	---	---	<0.001	<0.001	<0.001	---	---	---
	Control	Left position	0.203	<0.001	<0.001	---	---	---	<0.001	<0.001	<0.001	---	---	---
		Right position	0.299	<0.001	<0.001	---	---	---	<0.001	<0.001	<0.001	---	---	---
		Dog	0.958	0.483	0.172	0.565	0.417	0.149	<0.001	<0.001	<0.001	0.009	<0.001	<0.001
		Pair of people	0.097	0.752	0.018	0.476	0.139	0.271	<0.001	<0.001	0.001	<0.001	0.070	<0.001
		Vehicles	0.011	0.002	0.843	0.210	0.154	0.073	<0.001	<0.001	<0.001	<0.001	0.008	<0.001
	Obscuring face & body	Obscured face	0.017	0.089	0.224	---	---	---	<0.001	<0.001	<0.001	---	---	---
		Obscured body	0.020	0.112	0.238	---	---	---	<0.001	<0.001	<0.001	---	---	---
Block 2	**Primary**	-------	0.017	0.233	0.152	---	---	---	<0.001	<0.001	<0.001	---	---	---
	Control	Left position	0.746	0.005	0.001	---	---	---	<0.001	<0.001	<0.001	---	---	---
		Right position	0.012	0.014	0.007	---	---	---	0.006	0.006	0.004	---	---	---
		Dog	0.006	0.572	0.062	<0.001	0.334	0.006	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
		Pair of people	0.010	0.068	0.407	0.996	0.004	0.040	<0.001	<0.001	<0.001	<0.001	0.001	<0.001
		Vehicles	0.020	0.011	0.986	0.027	0.040	0.005	0.007	0.007	0.001	0.007	0.003	0.002
	Obscuring face & body	Obscured face	0.014	0.260	0.109	---	---	---	<0.001	<0.001	<0.001	---	---	---
		Obscured body	<0.001	0.026	0.000	---	---	---	<0.001	<0.001	<0.001	---	---	---

* Background
** Distraction(s), this could be either dogs, people, or vehicles, as shown in the subcategory.
Shaded= significant using Bonferroni corrections.

7.3.2. First fixations

The previous section of the analysis was reported the duration of fixations that fell on the test images. For further eye-tracking analysis, this section described on the first fixations that happened on the test images, which is a common measure of interest in eye-tracking research (Holmqvist et al., 2011, p.112). The first fixation is defined as one that occurs after the first saccade made at the start of the trial. Thus, it indicates the item that received the most visual attention rather than which item received the highest amount of viewing time (duration of fixations) as measured by the fixation durations.

Table 7.5 presents the results of first fixations (as percentages) across male and female targets for each domain in the test images applied to the normal-size and area-weighted scores in Blocks 1 and 2. Next part will discuss separately these results for each block.

In Block 1, Table 7.5 (a) shows the results of the normal-size. The highest percentage of first fixations in the primary images was towards the face (55%), followed by the background (30%), and finally the body (15%). This indicates that first fixations were higher towards the face when the target person was in the middle of the scene.

In the control images where the target person was not in the middle of the scene, the highest percentage of first fixations was towards the background (51% when the person was on the left, 56% on the right); with reduced attention to the face (17%) when target was on the left, or the body/background (22%) when the target was on the right. This indicates first fixations towards faces reduced when the target person was not in the middle of the scene. In subsequent control images where the target person was in the middle but the scene included a distraction item(s), when the distraction was a dog, the highest percentage of first fixations was on the background (49%), followed the face (29%), and finally the dog (5%). When the distraction was a pair of people, the highest percentage was on the face (43%), followed the background (25%), and finally the body (19%). When the distraction was a vehicle(s), the highest percentage was on the vehicle(s) (37%), followed by the face (36%), and finally the background (12%). Among the distraction items used, the vehicles were the most distracting, but when there was a pair of people in the scene, first fixations toward the person (especially the face) were higher than when vehicles or dogs were included.

In the obscuring face & body images, the highest percentage of first fixations was on the face (44% obscured face, 46% obscured body), followed by the background when the face was obscured (26%), or body (36%) when the body was obscured, and finally the body when the face was obscured (26%), or background (16%) when the body was obscured. Across all test images, the greatest percentage of first fixations was towards the face (37%), followed by the background (33%), the body (22%), and finally the distraction (8%).

The percentages of first fixations were then weighted by area shown in Table 7.5 (b). For all domains, the highest percentage was on the face (47% – 97%), followed by the body (2% – 50%), the distraction if there was one (10% dog, 8% people, 3% vehicles), and finally the background (0.2% – 5%). The control item that affected the first fixations towards the face was the left and right positions, where percentage on the faces was reduced to the lowest levels (47% left, 48% right). Across all test images, the greatest percentage was towards the face (81%), followed by the body (15%), and finally the background and distractions (1%).

In Block 2, where there was a question after each image, the results were different from the Block 1 where there was no question. Only the images of the obscured body yielded results similar to Block 1. In the normal-size in Table 7.5 (a), the images yielded substantial changes, although in primary images, the distribution of first fixations was similar with low differences (4%, 2%). The highest percentage changed from the face to the body, with the percentage on the face falling from 55% to 32%, while the percentages in the body and background increased from 15% to 36% (body) and 30% to 33% (background). This indicates that the effect of thinking about a safety question suggests more observation of the scene where the percentages of first fixations were very similar between test domains and no dominant item was found.

In the control images where the target person was not in the middle of the scene, the highest percentage of first fixations changed from the background to the face; in the left position, first fixations on the face increased from 17% to 53%, while first fixations on body and background decreased from 32% to 27% , and from 51% to 21%, respectively. For the right position, first fixations on the face increased from 22% to 56%, while first fixations on the body did not change (22%), and first fixations on the background decreased from 56% to 22%.

In the next control image where the target person was in the middle of the scene, and a distraction item(s) was included, when the distraction was a dog, the highest percentage of first fixations changed from background to face; with face increasing from 29% to 51%, and body increasing from 18% to 25%, while background decreased from 49% to 25%, and dog decreased from 5% to 0%. When the distraction was a pair of people, the highest percentage of first fixations continued to be on the face, but this decreased slightly from 43% to 40%; fixations on body increased from 13% to 28%, and on people from 19% to 22%, while fixations on background decreased from 25% to 11%. When the distraction was vehicles, the highest percentage of first fixations changed from vehicles to background; fixations on face decreased from 36% to 31%, and vehicles from 37% to 6%, while fixations on body increased from 15% to 27% (body), and from 12% to 35% (background).

In the obscuring face & body images, when the face was obscured, the highest percentage of first fixations changed from obscured face to body; fixations on face decreased from 44% to 33%, and on background from 29% to 25%, while fixations on body increased from 26% to 42%. By contrast, when the body was obscured, the highest percentage continued to be on the face; fixations on face increased from 46% to 51%, and on background from 16% to 20%, while fixations on obscured body decreased from 38% to 29%. Across all test images, the results were similar to Block 1: the greatest percentage of first fixations among the test domains was towards the face (increased from 37% to 43%); it also increased on the body from 22% to 29%, while it was decreased on the background from 33% to 24%, and on the distractions from 8% to 4%.

The percentages of first fixations were then weighted by area shown in Table 7.5 (b). For all domains, the highest percentage was on the face (85% – 96%), followed by the body (4% – 10%) if there were no distractions, or distractions (0% dog, 9% people, 0.5% vehicles) if these were present, and finally the background (0.1% – 0.6%). The control item that affected the first fixations towards the face was the body, where the percentage on faces reduced to their lowest percentage (47%), and the percentage on the body was the highest (50%). Across all test images, the greatest percentage of first fixations was towards the face (92%), followed by the body (6%), distractions (1%), and finally the background (0.3%).

Overall, the highest percentage of fixation durations and first fixations in the normal-size was towards the face, and it was significantly increased in the area-weighted. However, one limitation in this analysis concerns the effect of the number of images as these were not equal due to the randomisation process.

Similar to the analysis of the duration of fixations, the first fixations in this experiment were analysed regardless of age and gender or other characteristics. Results for each gender and block that show first fixations (percentage associated with number of images) are reported in Appendix N, Tables N.7 and N.8.

Table 7.5. Percentage of the first fixations along with the number of images across male and female targets in each domain of the test image in Blocks 1 and 2. *(a)* The scores of the normal-size. *(b)* The scores of the area-weighted.

Block	**Category of test images**	**Subcategory**	**(a) Normal-size**								**(b) Area-weighted**							
			Test domains on the image:								**Test domains on the image:**							
			Face		**Body**		**Background**		**Distraction**		**Face**		**Body**		**Background**		**Distraction**	
			%	**n***	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**
Block 1	**Primary**	-------	55%	11	15%	13	30%	8	---	---	97%	11	2%	13	0.2%	8	---	---
	Control	Left position	17%	2	32%	20	51%	26	---	---	47%	2	50%	20	3%	26	---	---
		Right position	22%	11	22%	3	56%	33	---	---	48%	11	46%	3	5%	33	---	---
		Dog	29%	7	18%	15	49%	9	5%	1	84%	7	5%	15	1%	9	10%	1
		Pair of people	43%	7	13%	10	25%	12	19%	3	89%	7	2%	10	0.2%	12	8%	3
		Vehicles	36%	9	15%	12	12%	5	37%	6	93%	9	4%	12	0.2%	5	3%	6
	Obscuring face & body	Obscured face	44%	10	26%	21	29%	16	---	---	93%	10	6%	21	0%	16	---	---
		Obscured body	46%	14	38%	12	16%	22	---	---	93%	14	7%	12	0.1%	22	---	---
	All images	-------	37%	71	22%	106	33%	131	8%	10	81%	71	15%	106	1%	131	1%	10
Block 2	**Primary**	-------	32%	19	36%	9	33%	4	---	---	90%	19	10%	9	0.4%	4	---	---
	Control	Left position	53%	6	27%	18	20%	17	---	---	95%	6	4%	18	0.2%	17	---	---
		Right position	56%	11	22%	9	22%	32	---	---	96%	11	4%	9	0.2%	32	---	---
		Dog	51%	13	25%	12	25%	7	0%	0	96%	13	4%	12	0.2%	7	0%	0
		Pair of people	40%	11	28%	9	11%	4	22%	7	85%	11	5%	9	0.1%	4	9%	7
		Vehicles	31%	16	27%	7	35%	6	6%	3	92%	16	7%	7	0.6%	6	0.5%	3
	Obscuring face & body	Obscured face	33%	8	42%	32	25%	7	---	---	89%	8	10%	32	0.3%	7	---	---
		Obscured body	51%	25	29%	6	20%	14	---	---	95%	25	5%	6	0.2%	14	---	---
	All images	-------	43%	109	29%	102	24%	91	4%	10	92%	109	6%	102	0.3%	91	1%	10

7.4. Summary

This chapter presented the results of the eye-tracking – Experiment 3 that was conducted to provide an objective measure of the visual evaluation of other pedestrians by repeating the study undertaken by Fletcher-Watson et al. (2008). A key objective was to overcome the limitations of their work and apply validation steps, and to answer the question: is the face an important visual cue?

The analysis of fixation durations on a series of test images that compared the duration times of each fixation suggested that people tend to look at the face of a person. The analysis of the primary images where the target person was in the middle of the street revealed a significantly greater fixation of durations towards the face than the body and the background. Several questions were then addressed, such as: Does it matter if the target was in the centre or at the side of the scene? Does it matter if there were distractions in the scene such as dog, vehicles, and other people? Direct comparisons between these items revealed significantly greater mean duration on the face. This suggested people tend to look at the face of a person more than other items in the scene. This conclusion was the same in the two test blocks; where Block 1 was involved a free viewing with no other task, and Block 2 required a response to a question after each image. The general comparison of fixation durations between Blocks 1 and 2 suggested a significant difference, which indicates that gaze behaviour were affected by the safety question on people's minds; same of this conclusion was found in the first fixations analysis

Analysis of the first fixations of normal-size in Block 1 revealed that 37% of first fixations fell on the face, and this increased in the area-weighted to 81%; similar results were found in Block 2 where 43% of first fixations fall on the face, increased the area-weighted to 92%. The lowest percentage of fixations of the normal-size in Block 1 fell on the distractions with 8%, decreasing in Block 2 to 4%. However, the results in the area-weighted images differed as the percentage of fixation durations on the background decreased from 33% to 1% in Block 1, and 24% to 0.3% in Block 2; this was expected as the large size of the background was reduced when it was weighted. The analysis also indicates that even when the face or the body is obscured, a higher fixation duration and number of first fixations were found towards the face. Consequently, we can answer the raised question by yes, the face is an important visual cue. The next chapter presents an overall discussion of this thesis, including the limitations of the experiments that were conducted.

Chapter 8. Discussion

8.1. Introduction

The aim of this work was to determine whether the face of an approaching pedestrian is a critical visual cue in interpersonal evaluations. If it is, then previous research on lighting and face-based evaluations (of either identity or intent) can be assumed to be relevant. If it is not, then this body of research becomes irrelevant for its intended purpose. Three experiments were performed to test the two research hypotheses:

H1: The face as a visual cue has an influence on interpersonal evaluations.

H2: The face is a more important visual cue in interpersonal evaluations than the body or the hands.

Experiments 1 and 2 (Chapters 3, 4 and 5) employed similar procedure which was to elicit evaluations of safety when observing images of actors portraying pedestrians in different situations. The core assumption was that those targets evaluated as least safe would be those obscuring the more important visual cue(s). However, these subjective evaluations are subject to numerous forms of unintended bias (Poulton, 1977). Therefore, Experiment 3 (Chapters 6 and 7) was conducted using eye-tracking, an objective measure, and was based on the assumption that the more important visual cues in a scene are those which receive the first visual fixation and the longest duration of visual fixations. The current chapter assesses whether the results of these experiments support or refute the hypotheses proposed. It then discusses whether these findings agree or disagree with previous work, the limitations of the current work, what the findings mean for lighting practice, and makes suggestions for further research.

8.2. The important visual cues for pedestrians

Experiment 1 explored the influence of five visual cues (gender, number of people, walking direction, light direction, and the exposure of face and hands) of approaching pedestrians. The cues were tested using 16 test images of target person(s) portrayed the cues, and it presented on a PC screen. Table 8.1 presents the differences between mean/median ratings for each cue which were analysed to identify a trend in what were perceived as less safe conditions; the table also presents statistically significant differences for each cue that was tested in pairs. The results demonstrated that the exposure of face and hands had a greater effect than the other tested cues, yielding the greatest difference between mean/median ratings and a higher number of statistically significant pairs between test images. This result supported H1 which stated that the face has an influence on interpersonal evaluations.

Table 8.1. Results for the predicted less safe conditions and the significant differences for each visual cue.

Visual cues	Predicted less safe	Difference between		Trend supports prediction*	Number of pairs suggested to be significant**	Overall decision for the significant difference
		Median	Mean			
Gender	Female	0	0.3	Yes	3 of 6	Yes
Number of people	One person	0	0.2	No	2 of 3	Yes
Walking direction	Away	0	0.2	Yes	3 of 6	Yes
Light direction	Front	0.5	0.3	Yes	5 of 6	Yes
Exposure of face and hands	Exposed	1	0.6	Yes	5 of 6	Yes

* Based on the difference between the mean/median ratings.
** Based on the differences between pairs for each cue analysed using the Wilcoxon test (see Sections 4.3.1 to 4.3.5).

Experiment 2 used the experiences and findings from Experiment 1 to test H2 and conduct further investigation into the influence of visual cues. The experiment focused on face and hands which were found to have a greater influence on safety than other cues. Experiment 2 also addressed the limitations found in Experiment 1 (discussed in Section 3.6), and certain steps were built in to enhance content validity; for instance, participants were asked to walk on a treadmill, and a recording of street sounds was played during trials. Two procedures of evaluation were applied: category rating and paired comparisons – to increase the robustness of the results. For each procedure, additional images were included to check validity: repeated images in category rating, and null condition trials in paired comparisons (see Sections 3.4.1 and 3.4.2, respectively).

The analysis compared safety ratings for different levels of exposure of face and hands across three test blocks (face block, hands block, and mixed block – face and hands). The findings from category rating data indicated that evaluations of safety were reduced as the face and hands became more concealed, and changes in face concealment led to greater evaluations than changes in hand concealment (Figure 8.1). This suggests that hands behind the body are viewed as less safe than other hand positions (at the side of the body, in front of the body, and inside the pockets); the results of the mixed block were similar, and supported the findings of the face and hands blocks (see Section 5.2.1). This conclusion was also supported by the paired comparison data which exhibited similar trends, the exceptions being that hands in front of the body and inside the pockets were rated lower (Figure 8.2). The statistically significant differences between poses for each block in category rating and paired comparisons procedures supported this (see Sections 5.2.1 and 5.3.1, respectively). These findings supported H2 which stated that seeing the face is more important than seeing the hands for interpersonal evaluation.

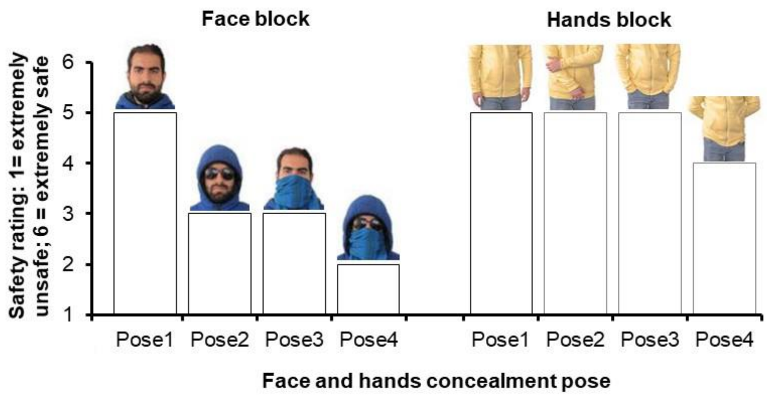


Figure 8.1. Comparison between the results of face and hands blocks for the category rating procedure which shows the median safety rating for each pose across all actors.

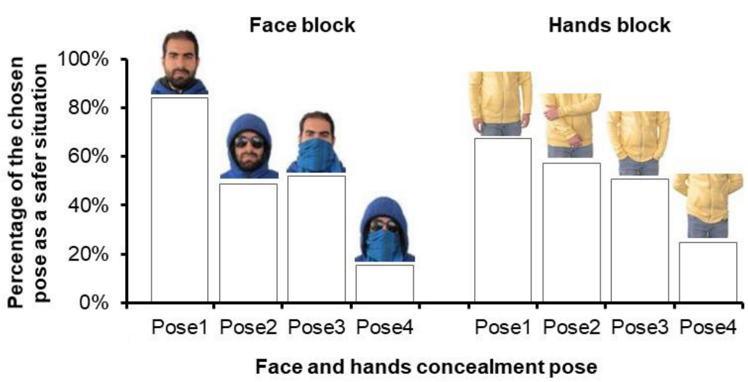


Figure 8.2. Comparison between the results of face and hands blocks for the paired comparisons procedure which shows the percentage of participants choosing the safer pose.

Moreover, two background scenes were used in Experiment 2 to determine whether these influenced the safety evaluations. The results from both procedures did not suggest an effect of the background scene, with no difference found in the interpersonal evaluation of other pedestrians between a brightly-lit road and dimly-lit back alley (see Sections 5.2 and 5.3).

The findings from Experiments 1 and 2 were further verified by Experiment 3 which employed an entirely different approach (eye-tracking). This experiment repeated an experiment by Fletcher-Watson et al. (2008), but overcame a key limitation by including validation steps (see Section 6.1). Test participants observed a set of images, single image presented consecutively, some of which contained a person and some of which did not. The analysis targeted only the images containing a person to answer the question: is the face an important visual cue? This involved comparing durations of fixations and locations of first fixations towards three test domains: face, body, and background, along with distraction items such as changing the position of the target person and adding other domains into the scene, such as other people, vehicles, and dogs.

The results of Blocks 1 and 2 indicated that the duration of fixations and first fixations was higher towards faces than other items; however, these fixations were reduced when the target shifted from the middle (moving to the left or right side), resulting in the background receiving a higher level of first fixations and duration of fixations. Conversely, although the inclusion of distractions in the scene, faces continued to receive the most first fixations and greater durations of fixations (Table 8.2 (a)). These findings were then weighted to give relative measures between domains by applying an area-weighted analysis (see Section 6.3). The results revealed greater durations of fixations and first fixations towards the face more than all other domains (Table 8.2 (b)).

In summary, the findings of Experiment 3 suggest people tend to look at the face of a person more than the body, and more than other items in the scene such as dogs, vehicles, and other people walking away. Overall, the results of all these experiments support the assumption that the face is an important visual cue, and hence a valid indicator for the interpersonal evaluation of pedestrians.

Experiment 3 included some images in which the actors body and face were exposed and some in which the body or face were obscured. Consider those images where the body was obscured so it did not provide a cue to the location of the face.

Results of duration of fixations from Block 1 (free viewing) for normal-size images (Table 8.2) show that in face present images 41% of fixations by duration were towards the face, increasing to 43% when the body was obscured. This confirms a tendency to look at the face. When the face was obscured to leave only the body, 23% of fixations by duration were towards the space where the face would be expected.

Similarly, results of first fixations from Block 1 (free viewing) for normal-size images (Table 8.2) show that in face present images 55% of first fixations were towards the face, decreasing to 46% when the body was obscured. When the face was obscured to leave only the body, 44% of first fixations were towards the space where the face would be expected.

This confirms the same tendency to look at the face found in the duration of fixations. This suggests an expectation to see a face above the body, but not by a sufficient magnitude to explain all fixations towards the face. Similar results were found in the Block 2. Overall, this does not show a strong similarity in gaze behaviour towards the face, but it was found in the area-weighted analysis (Table 8.2).

Table 8.2. Results of fixation durations and first fixations (as percentages) across male and female targets in each domain in normal-size and area-weighted scores for Blocks 1 and 2.

Block	Category of test images	Subcategory	Duration of fixations								First fixations							
			(a) Normal-size				(b) Area-weighted				(a) Normal-size				(b) Area-weighted			
			Face	Body	Bac*	Dis**	Face	Body	Bac*	Dis**	Face	Body	Bac*	Dis**	Face	Body	Bac*	Dis**
Block 1	**Primary**	-------	41	26	33	---	94	5	0.3	---	55	15	30	---	97	2	0.2	---
	Control	Left position	19	26	56	---	88	11	1	---	17	32	51	---	47	50	3	---
		Right position	21	16	63	---	92	6	1	---	22	22	56	---	48	46	5	---
		Dog	25	24	30	21	69	6	0.4	25	29	18	49	5	84	5	1	10
		Pair of people	29	16	31	24	88	4	0.4	7	43	13	25	19	89	2	0.2	8
		Vehicles	38	18	17	27	94	4	0.2	1	36	15	12	37	93	4	0.2	3
	Obscuring face & body	Obscured face	23	42	34	---	86	14	1	---	44	26	29	---	93	6	0.4	---
		Obscured body	43	24	32	---	95	5	0.3	---	46	38	16	---	93	7	0.1	---
	All images	-------	25	21	33	21	82	6	0.5	11	37	22	33	8	81	15	1	3
Block 2	**Primary**	-------	45	22	34	---	95	4	0.3	---	32	36	33	---	90	10	0.4	---
	Control	Left position	27	25	47	---	92	8	1	---	53	27	20	---	95	4	0.2	---
		Right position	33	13	54	---	96	3	1	---	56	22	22	---	96	4	0.2	---
		Dog	36	18	32	15	81	4	0.3	15	51	25	25	0	96	4	0.2	0
		Pair of people	32	16	21	32	87	4	0.3	9	40	28	11	22	85	5	0.1	9
		Vehicles	46	15	15	25	96	3	0.2	1	31	27	35	6	92	7	1	0
	Obscuring face & body	Obscured face	26	42	32	---	87	13	0.5	---	33	42	25	---	89	10	0.3	---
		Obscured body	51	13	35	---	97	2	0.3	---	51	29	20	---	95	5	0.2	---
	All images	-------	31	17	31	21	88	4	0.4	7	43	29	24	4	92	6	0.3	1

* Background
** Distraction

8.3. Limitations

In real-life situations, pedestrians encounter and evaluate other pedestrians when walking along roads. In this work, the evaluations were performed in a laboratory while observing images of pedestrians. Thus, safety evaluation in a laboratory is not natural – it captures imagined safety. The differences in these settings include the degree of insecurity imposed by the actors and the environmental, and image characteristics of the observed scenes. Consequently, the content validity was reduced, and the test images and the laboratory setup were not expected to convey the same sense of fear/safety as would be the case when encountering real pedestrians.

One of the main reasons for using images and the laboratory environment was the cost and the repeatability of the experiment. To illustrate, paying real actors and participants to contribute to an experiment for several days would incur high expenses, and it is not guaranteed that the same actor will be present and able to repeatedly make the same facial expressions and poses.

Other reason for using images instead of real persons was to exclude unrelated changeable variables that might influence the results such as light conditions, time, and elements in the background such as road familiarity and building types. Nevertheless, two limitations of using images were noted:

1. The results may have been influenced by the nature of the images and/or the actors presented to participants in trials. For instance, test participants may have reported the concealed face as less safe than the exposed face because they thought that was the expected response. This is an example of participant bias, where participants adjust their behaviour in relation to what they perceive to be the experimenter's expectations (Orne, 1962).
2. The findings may be influenced by the experimenter-induced stimulus range bias where the experimenter’s choice of images might has a direct bearing on the participants’ responses (Ward, 1972).

It may be thought that conducting those experiments using real actors in real-life situations would provide more relevant responses than using images of actors observed in a laboratory environment. However, a study by Hariyadi & Fukuda (2017) suggests otherwise. They conducted two experiments to compare evaluations of the visibility of objects behind window blinds; one experiment used digital images and the other used real conditions. They found very similar responses between the two experiments which suggests that evaluations based on images are a sufficient proxy for real-world conditions. The same conclusion was drawn in a second study, in which differences between subjective evaluations of real and virtual environments were not reported to be significantly different (Chamilothori et al., 2018). Nevertheless, more research is needed in order to determine whether the findings made in this study can be generalized to real-life outdoor settings.

Another reason for conducting a laboratory experiment instead of a real-life situation was to control test variables precisely and reduce variations in fear. For instance, interpersonal evaluation in a field may be influenced by surrounding elements such as other people and building types (e.g., grocery store, police station, or a pub).

The test environment in the laboratory was set up to promote a level of visual adaptation similar to that of an outdoor situation, and to target approximately representative levels of luminance adaptation; however, the light measurements were found brighter than real outdoor space after-dark. Table 8.3 details the vertical illuminance at the observer's eyes, and the horizontal illuminance for the experiments. The former was 4 – 23 lx which is considered brighter than reality, as a good road lighting situation would typically be 10 lux (Boyce, 2014, p. 9, Table 1.3). This was due to the large size of the projected image on the wall and/or the brightness intensity of the used background. However, a change in light level was not expected to significantly affect relative responses for the different face and hands poses. Nevertheless, further research is needed to determine the extent to which the evaluations made can be generalised to those made in real outdoor settings.

Table 8.3. Light measurement results for the experiments conducted.

Illuminance (lux)	**Light measurement in**			
	Experiment 1*	**Experiment 2* (Scene 1)**	**Experiment 2* (Scene 2)**	**Experiment 3***
Vertical	12	18	4	23
Horizontal	14	21	5	25

* Figures 3.12, 3.13, and 6.19 shows measurement locations for Experiments 1, 2, and 3, respectively.

The results of Experiment 3 suggested a tendency to look at the faces of other people. One potential limitation is that the faces of the target people were all at about the same height – about the middle, the eye line of the test participants. That does, however, represent the natural situation: faces (of adults) tend to be at around the same height; ranging between the shortest at 1.60 m (Timor-Leste) and the tallest at 1.82 m (Netherlands) of adult men (https://ourworldindata.org/human-height). So, if a pedestrian is walking and looking straight forward, the faces of others will be in the region represented in the images.

Controls were included in this experiment to check for the influence of automatic gaze behaviour. This would check, for example, whether participants looked at the face because it was a face or because of a natural tendency to look at the middle of the screen. To investigate horizontal scanning, the actors' position was changed from the centre to either on the left and right sides of the scene. This was on the same horizontal line to represent the expected similar heights of adult faces above ground level (Figure 8.3). It would be unrealistic to expect people at a higher or lower level (Figure 8.4) unless they were scaled in size to represent different interpersonal distances (as shown in Figure 8.5) and this could confound the analyses of fixations on different objects and locations. In further work it would be interesting to include images such as Figure 8.5 to investigate the impact of different interpersonal distances and object sizes.

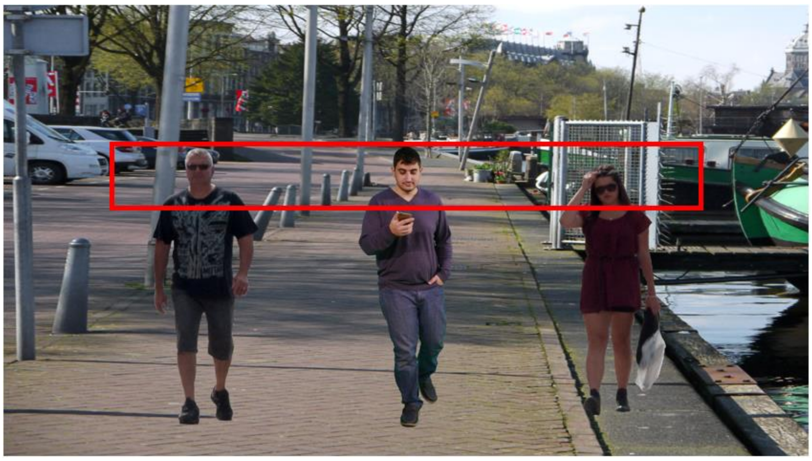


Figure 8.3. The three positions of the actors at 4 m in the same horizonal line that used separately in Experiment 3.

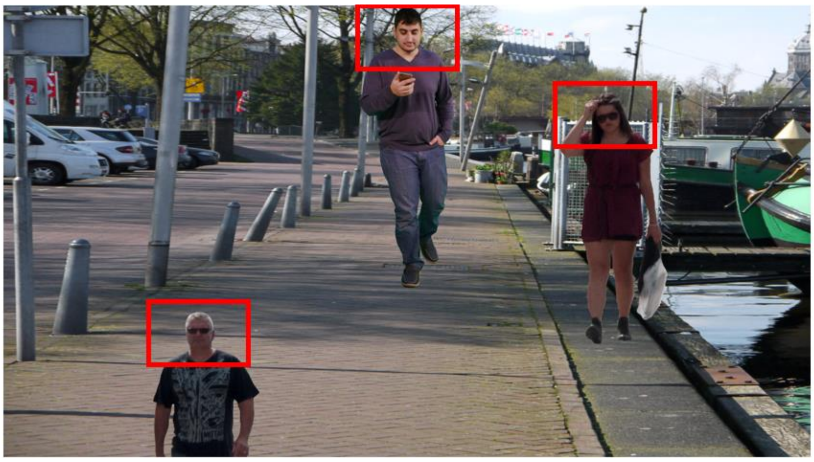


Figure 8.4. Example of changing the positions of the actors in a vertical line, while keep them at the same interpersonal distance (4 m).

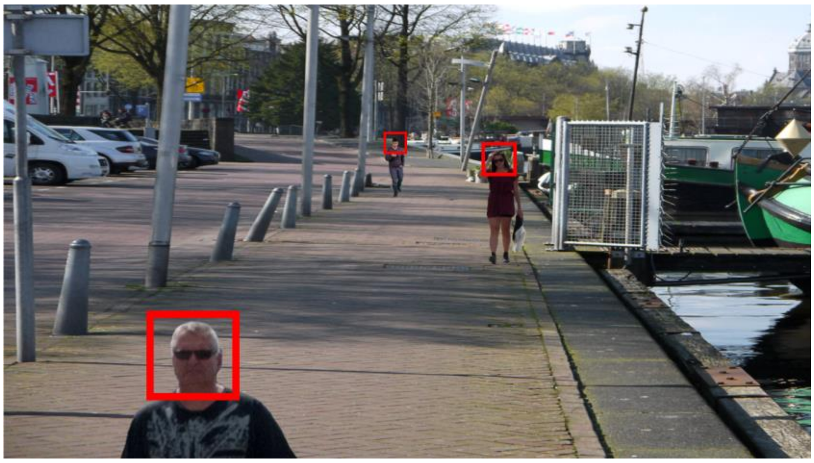


Figure 8.5. Example of changing the positions and the size (i.e. distance) of the actors in a vertical line, representing interpersonal distances of 2 m (left), 10 m (right), and 15 m (middle).

Because of these limitations, this thesis conducted three experiments using different methods so that limitations in one method could be overcome by the use of another. Thus, the same conclusion drawn from multiple methods tends to be robust and reliable. Future studies could explore whether evaluating other people in a real street differs from evaluating them in the laboratory, and considering the diverse characteristics of actors such as ethnicities, familiarity, and ages.

Experiments 1 and 2 evaluated the safety ratings of other pedestrians by asking the participants: how safe would you feel in this situation? In previous lighting studies, investigating interpersonal evaluations based on FIR and FER in real and laboratory experiments using real persons or images (Table 8.4). Participants in these studies were not asked about safety, they were asked about recognisability, identification, or visibility. The results of the light levels and SPD effect in these experiments were mixed (see Table 2.8). One possible reason is that each study has tended to use unique conditions (see Tables 2.6, and 2.7), and/or the researchers' degrees of freedom might have led to them making biased decisions (Wicherts et al., 2016). Therefore, a different approach to investigating the effect of changes in lighting on interpersonal evaluations is to use a model of visual performance focusing on relative visual performance (RVP).

The RVP model is an attempt to characterise the visual component of visual task performance by minimising the cognitive and motor elements of task performance. It does this by using a task in which the visual part is large and the motor and cognitive parts are small, and measures the speed and accuracy with which task is carried out (Rea & Ouellette, 1991). One such task was numerical verification, comparing two columns of 5-digit of numbers, as developed by Smith & Rea (1980). RVP provides a measure of the effect of lighting on performance of the visual part of a task, and by characterising the task by contrast and size (alongside the adaptation luminance and the observer age) can be applied to any task (Boyce, 2014, p. 136).

The ability to evaluate different facial expressions or facial identities can be considered as the ability to recognise or discriminate between a series of luminance patterns created by facial features. RVP should then be able to characterise the degree to which facial features 'could' be distinguished, albeit without making any direct evaluations of facial identity or facial expression. This is similar to an optician checking eye sight using a Snellen chart (individual letters) not words or sentences. If a certain situation creates higher RVP then it would be expected that the ability to make face-based evaluations would also be increased. It predicted whether we can see detail of, say, the mouth, but not an observer's assessment of mouth shape. Thus, RVP is an alternative to experimental studies of facial identification and facial emotion. By characterising the change in RVP with different adaptation luminances, RVP provides a prediction of the optimal lighting for a task of given contrast, size, and observer age. What should then happen is confirmation of that prediction by experiment. One limitation of RVP is that it considers only the adaptation luminance, and does not account for other changes in lighting such as SPD.

Hence, predictions of visual performance using the RVP model was recently used by Fotios et al. (2022), they found that such an evaluation is affected by an evaluation of this type is affected by the adaptation luminance, reflectance of the pavement surface, observer age, and skin tone of the observed person.

Table 8.4. Examples of past studies in lighting focusing on FIR and FER that were conducted in the field or in a context of outdoor lighting (laboratory).

Carried out in	Nature of target person	Studies of interpersonal evaluation based on	
		FIR	**FER**
Field	**Real**	Alferdinck et al. (2010), Boyce & Rea (1990), Rea et al. (2009), Rombauts et al. (1989), Johansson et al. (2001)	---------
	Image	Lin & Fotios (2015), Knight et al. (2007), Knight (2010), Romnée & Bodart (2014), Yao et al. (2009)	---------
Lab	**Real**	Raynham & Saksvikronning (2003), Okud & Satoh (2000), Caminada & Van Bommel (1984), Iwata et al. (2015)	---------
	Image	Dong et al. (2015)	Fotios, Yang & Cheal (2015), Yang & Fotios (2015), Fotios et al. (2017), Johansson & Rahm (2015), Fotios et al. (2018), Rahm & Johansson (2018)

8.4. Summary

This chapter compared the findings of the three conducted experiments with previous studies investigating interpersonal evaluation, which confirmed that focusing on the face is a reasonable assumption and validated those previous results. The limitations that found in the three experiments were discussed, which highlighted the need for further investigations into faces and interpersonal evaluations. One suggestion is to use a model of RVP to predict performance of a face-based interpersonal evaluation, and to link safety responses towards certain characteristics of people (i.e. participants) and/or the environmental. To illustrate, will female participants rate other females as safer than males, even if they are exhibiting the same threatening behaviour? If a participant was new to the road (or area), do they evaluate other people similarly to their familiar areas?

Chapter 9. Conclusion

9.1. Conclusions for this Work

This work investigated the visual cues used by pedestrians to evaluate other people (interpersonal evaluations). Its aim was to guide research, which in turn informs standards for pedestrian lighting after dark in subsidiary roads. Previous lighting research has tended to assume that the face is the most important visual cue, operationalised as recognition of facial identify or facial expression of emotion, but this assumption was drawn without any evidence or empirical basis. Other possible visual cues such as hand gestures, body posture, and gaze direction have received little or no attention in lighting research. However, the fact that they have different characteristics from the face (size and contrast of features, and type of information conveyed) suggests their use in lighting research might lead to different conclusions about optimal road lighting conditions. Following the literature review, two hypotheses were established:

> H1: The face as a visual cue has an influence on interpersonal evaluations.
>
> H2: The face is a more important visual cue in interpersonal evaluations than the body or the hands.

These were tested in three experiments. Experiments 1 and 2 sought subjective safety evaluations of pedestrians displaying different characteristics. Experiment 1 included those characteristics suggested by a lighting designer to inform her evaluations of safety: the results suggested that the face and hands were dominant cue; the used test images did not allow discrimination between the face and hands – both were either exposed or concealed. Thus, Experiment 2 repeated Experiment 1 but used images designed to distinguish between the face and the hands. The results suggested that faces were a more important visual cue than the hands. Nevertheless, there were limitations to this approach such as the influence of the nature of the images and/or the actors (see Section 8.3). To rectify this, Experiment 3 was conducted using a different method – eye-tracking. The results suggested that people tend to fixate at the face of a person more than the body and other items in the scene such as dogs and vehicles. Therefore, H1 and H2 were supported by these findings.

The conclusion of this work confirms the assumption of previous lighting studies investigating the degree to which road lighting supports evaluations of facial identity and facial emotions (Alferdinck et al., 2010; Boyce & Rea, 1990; Caminada & Van Bommel, 1984; Dong et al., 2015; Knight, 2010; Knight et al., 2007; Lin & Fotios, 2015; Okud & Satoh, 2000; Raynham & Saksvikronning, 2003; Rea et al., 2009; Rombauts et al., 1989; Romnée & Bodart, 2014; Yao et al., 2009). These studies did not know the most relevant (and important) visual cues, which can influence the measurements employed and how the task is operationalised, thereby impacting the results and conclusions. Given the findings of the current study, future research on lighting and interpersonal evaluations should use the face as the most important visual cue.

9.2. Further work

A real person in real situations will allow for an accurate evaluation compared with actors in images presented in laboratory settings. Because the experiments in this thesis did not involve real persons and situations, further work will be needed to confirm the results (see Section 8.3).

In addition, the current results were obtained from a young test sample (ranging in age from 18 to 54 years) in which most participants were aged between 25 and 34 (73%). The results should therefore be verified with groups who might respond differently to the young, such as the elderly and people with impaired mobility. Compared with the young, the elderly are more likely to have reduced visual function, less physical mobility, and as pedestrians are at greater risk of fatality from road traffic crashes, and express greater vulnerability to personal attack (Department of Transport, 2020; Lagrange & Ferraro, 1989; Oxley et al., 2004; Sparrow et al., 2002). Using a broader sample will also promotes data which better represent society. If, for example, the elderly place greater emphasis on the visibility of the hands for their safety evaluations than those who are younger, this might change the conclusion that the face is the more important visual cue.

Furthermore, Experiments 1 and 2 were completed before the wearing of face masks became a normal habit (at least in the UK) due to the COVID-19 pandemic. This raises the question as to whether safety ratings based on face evaluations have changed during or after the pandemic, as people become used to seeing the lower part of people's faces covered? However, this was not the main focus of this thesis. The variations in face covering were merely a means to examine whether people use the face as a safety cue. Further research could nevertheless explore whether the pandemic has affected interpersonal evaluations based on the exposure of faces.

Appendices

Appendix A. Photographs of after-dark scenes

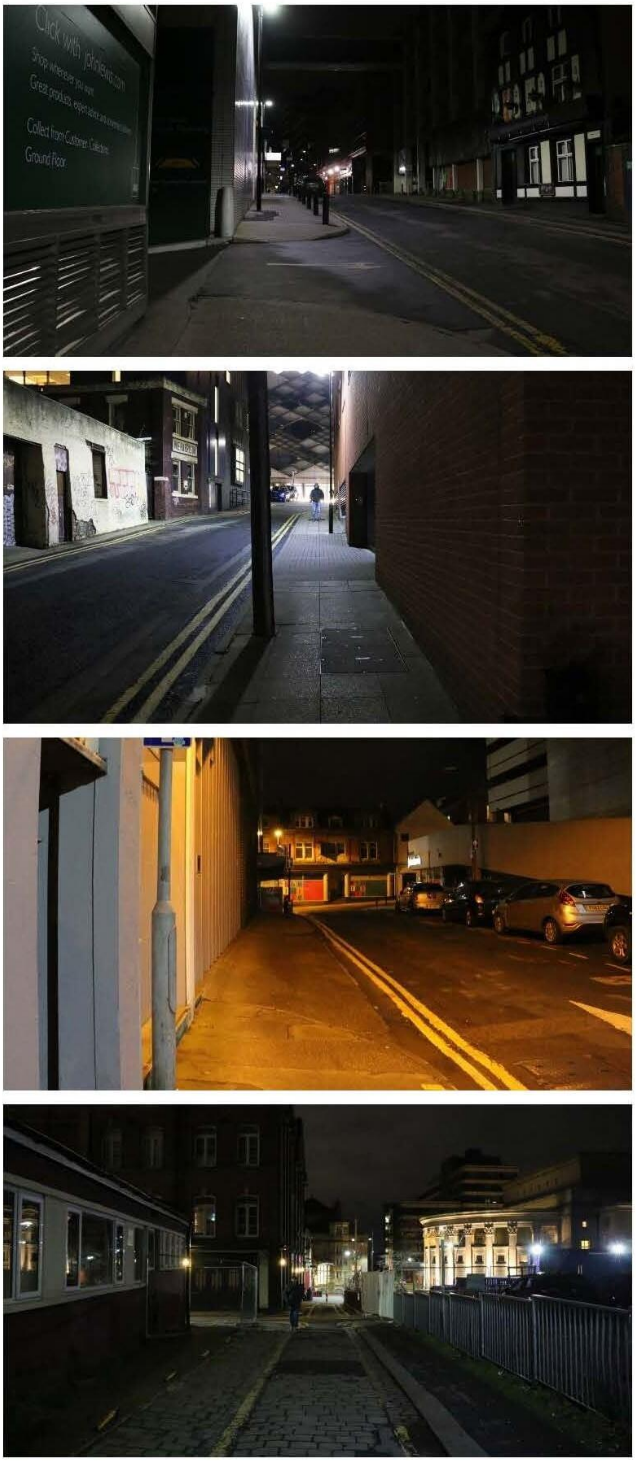


Figure A.1. Examples of different photograph scenes used to determine which the best to use for Experiments 1 and 2.

Appendix B. Methods of obscuring face and hands

Table B.1. Different methods for obscuring face and hands that were considered for Experiment 2.

Ways of obscuring	Face part			Hands
	Top	Bottom	Top + bottom	
Digital pixelating				
Digital black colour				
Hoodie or/and Mask				---------
Mask and/or sunglasses				---------

There were two studies that used different methods to obscure the face and body; Burton et al. (1999) used a black square, and Hahn et al. (2016) used a blurred square. Experiment 3 followed the same method as Hahn et al. (2016) with improvements to make it as realistic as possible (Table B.2). It was decided not to use the same obscuring method of those studies because replicating their method would offer no advantages, and our objective is to develop alternatives. In addition, the used obscuring method ensured no elements were added to the images (e.g. black square) that could influence the participants' responses. Therefore, the face and body were obscured by removing one of them and keeping a light shadow in its place, so it does not appear weird, and to aid the eye-tracking analysis by locating the boundary of the face and body.

Table B.2. Different methods for obscuring face and hands that were considered for Experiment 3.

Obscured part	**Studies**		
	Burton et al. (1999)	**Hahn et al. (2016)**	**Experiment 3**
Face	 Black square	 Blurred square	 Transparent
Body	 Black square	 Blurred square	 Transparent

Appendix C. Test images used for Experiments 1 and 2

Figure C.1. The 16 test images used in the Experiment 1.

Figure C.2. The 16 test images used in Experiment 2 for the face block, Scene 1. Note: images for Scene 2 are the same, but with a different background scene that was shown in Figure 3.2 *(b).*

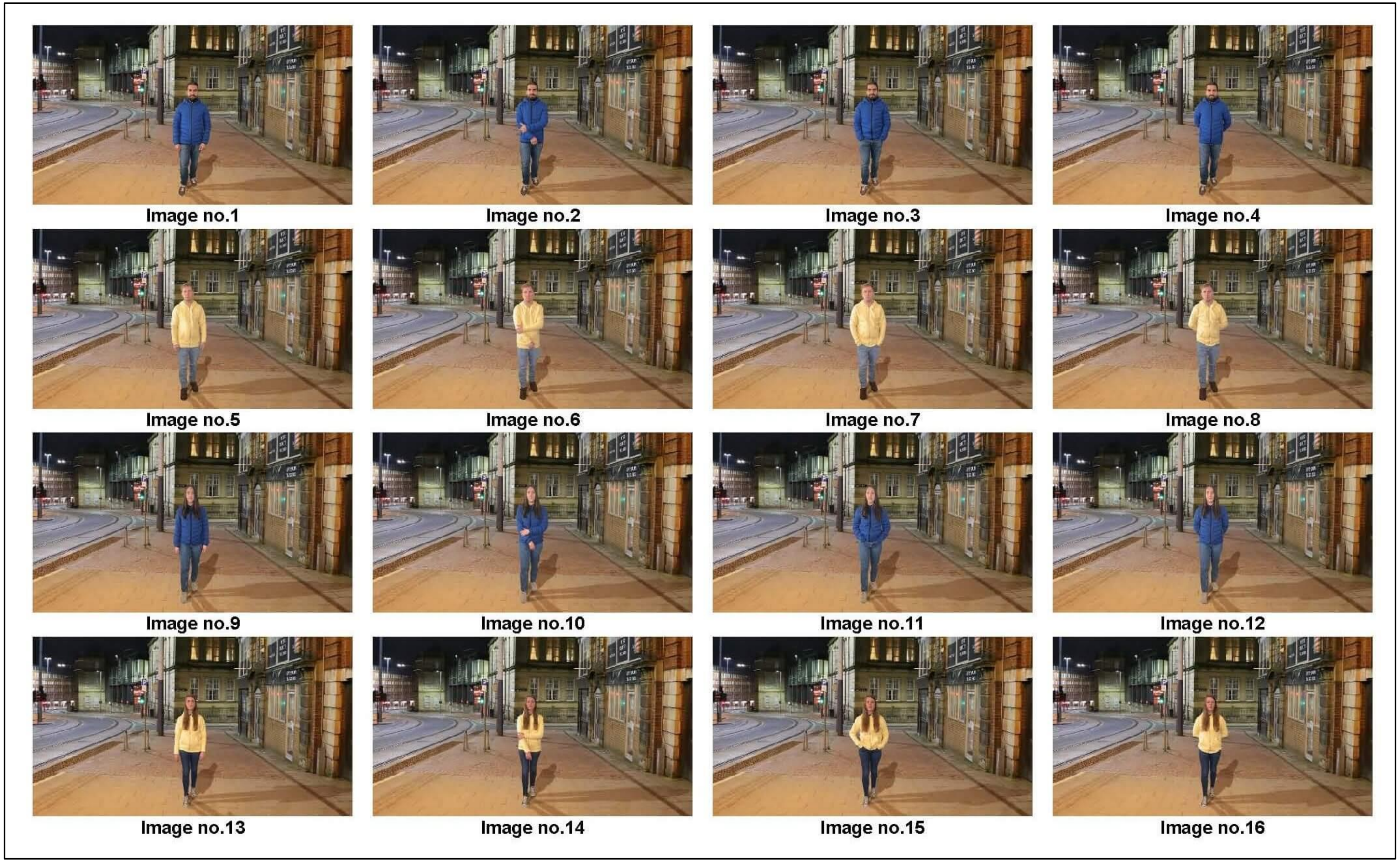


Figure C.3. The 16 test images used in Experiment 2 for the hands block, Scene 1. Note: images for Scene 2 are the same, but with a different background scene that was shown in Figure 3.2 *(b).*

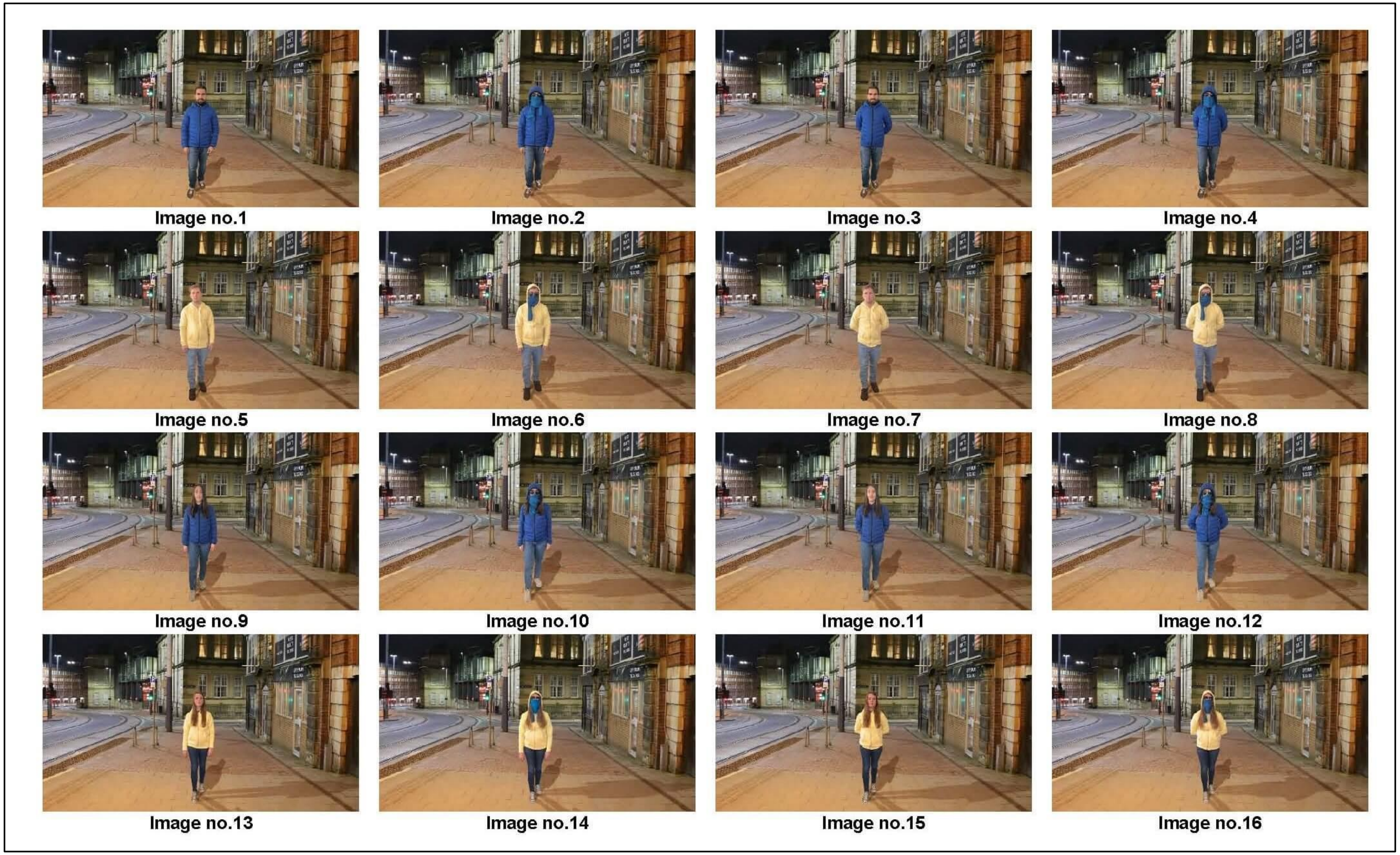


Figure C.4. The 16 test images used in Experiment 2 for mixed block, Scene 1. Note: images for Scene 2 are the same, but with a different background scene that was shown in Figure 3.2 *(b).*

Appendix D. Checklist form used in the experiments

Experiment 2 Checklist

Keys									
P.S	Play the sound	V.T.	Vision test	C.F	Consent form	D&P	Demo & practice	T.O	test order
H.S	Hands Single	H.P	Hands Paired	F.S	Face Single	F.P	Face Paired	M.S	Mix Single
M.P	Mix Paired	S.C	Scales comparison						

Test order:			
1	(H,F,M)	4	(F,H,M)
2	(H,M.F)	5	(M,F,H)
3	(F,M,H)	6	(M,H,F)
H: hands test		F: face test	
M: Mix test (hands and face)			

Age Groups	
1=18-24	5=55-64
2=25-34	6=65-74
3=35-44	7=75-84
4=45-54	8=85+years

Scene 1

				Adapt. time 20min				Main test										
Nu.	ID	G.	P.S	V.T.	C.F.	D&P	T.O.	H.S	H.P	F.S	F.P	M.S	M.P	S.C	Age group	speed	Break	Notes
1		M																
2		M																
3		M																
4		M																
5		M																
6		M																
7		M																
8		M																
9		M																
10		M																
11		M																
12		F																
13		F																
14		F																
15		F																
16		F																
17		F																
18		F																
19		F																
20		F																
21		F																
22		F																

11 M. 11 F.

Figure D.1. Checklist used in Experiment 2 for Scene 1; it used for the purpose of checking, randomisation, and other required tasks. Note: the checklist for Scene 2 followed the same form.

Experiment 3 Checklist and Randomization between blocks (Alternate way)

Sample size 32 people
First Block 1: 16 participants (8 male, 8 female)
First Block 2: 16 participants (8 male, 8 female)

Nu.	Gender	P. ID	Age	First block	Vision test	Consent form	Note
1	M			1	√		
2	M			2			
3	M			1			
4	M			2			
5	M			1			
6	M			2			
7	M			1			
8	M			2			
9	F			1			
10	F			2			
11	F			1			
12	F			2			
13	F			1			
14	F			2			
15	F			1			
16	F			2			
17	M			1			
18	M			2			
19	M			1			
20	M			2			
21	M			1			
22	M			2			
23	M			1			
24	M			2			
25	F			1			
26	F			2			
27	F			1			
28	F			2			
29	F			1			
30	F			2			
31	F			1			
32	F			2			

Figure D.2. Checklist used in Experiment 3; it used for the purpose of checking, randomisation, and writing notes.

Appendix E. Participant information sheets and consent forms used in the experiments

June 5th 2018

Participant Information Sheet

Using Lighting to Help Pedestrians Be Safe and Feel Safe

You are participating in a PhD research project that will be conducted at The University of Sheffield. Before deciding whether to take part it is important you understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Don't hesitate to ask if there is anything that is not clear or if you would like more information – please see contact details provided at the end. Take your time to decide whether or not your wish to participate.

1. What is the project's purpose?

In designing urban streets consideration must be given to how pedestrians' safety and their feelings of safety are continued after dark. The results of this experiment will contribute to ongoing research of current UK lighting standards for residential roads in order to give development recommendations for the British Standards Institution.

2. Why have I been chosen?

The study may have reached your email through the volunteer lists for staff or students maintained by CiCS. Other participants will be recruited from within staff and students of the University. The recruitment targets people not aware of the research, and to reach a semi-random sample.

3. Do I have to take part?

It is totally up to you to decide whether or not to take part. You can still withdraw at any time without any negative consequences, and you do not have to give a reason.

4. What will happen to me if I take part? What do I have to do?

The experiment will be conducted during normal working hours in the lighting laboratory on the 19th floor of the Arts Tower. The whole participation will not take more than 30 minutes. You will be given a copy of the information sheet and, if appropriate, a signed consent form to keep.

The project does not ask for any personal information except age and gender. Names will be strictly confidential and it will not be asked or mentioned in any phase of the research. All the information needed will be obtained from choosing pictures. In addition, participants will be asked to sign a consent sheet, and they have the right to withdraw at any time of the test without any harmful consequences.

The experiment will be conducted by showing on a PC screen a series of photographs of people in a street scene. The different people will all appear to be about 10 metres away from you. First, the photographs will be shown one at a time, each image will be visible for only 0.5 seconds. You will evaluate each image by answering the question "how safe do you

Figure E.1. Participant information sheet for Experiment 1.

June 5th 2018

feel in the vicinity of this person?" It is a multiple-choice answer on a five-point response scale from "very safe" to "very unsafe (you answer by clicking buttons underneath the images). This procedure will take about 2 minutes. The second procedure will show two images at the same time for 5 seconds, and you will choose the safer situation. This procedure asks for 120 comparisons, and it only needs 10 minutes to complete it.

5. What are the possible disadvantages and risks of taking part?

There are no known disadvantages to participation in this experiment. The researcher will be present and thus available to participants at all times if they need any help.

6. What are the possible benefits of taking part?

Participation in this experiment will help us to improve the standards for lighting design, which in turn leads toward a safe and protected community.

7. Will my taking part in this project be kept confidential?

All the information that we collect about you during the course of the research will be kept strictly confidential and will only be accessible to members of the research team. You will not be able to be identified in any reports or publications. If you agree to us sharing the information you provide with other researchers (e.g. by making it available in a data archive) then your personal details will not be included.

8. What is the legal basis for processing my personal data?

We will ask to record your age and gender and self-reported visual condition in order to demonstrate how well the data collected represent the whole population. We will not be recording your name or contact information.

9. What will happen to the data collected, and the results of the research project?

The experimenter will record each participants gender and age and self-reported visual condition. Each participant will be given a reference number. This means their responses are anonymised which give safety responses.

The data (anonymised test results) will be stored by the experimenter on their own PC with a back-up held by the supervisor.

Due to the nature of this research it is likely that other researchers may find the data collected to be useful in answering future research questions. We will ask for your explicit consent for your data to be shared in this way.

10. Who is organising and funding the research?

The researcher is funded by the Saudi Arabian Cultural Bureau.
The research is organized by The University of Sheffield, School of Architecture.

Figure E.1. (continued). Participant information sheet for Experiment 1.

June 5th 2018

11. Who is the Data Controller?

The University of Sheffield will act as the Data Controller for this study. This means that the University of Sheffield is responsible for looking after your information and using it properly.

12. Who has ethically reviewed the project?

This project has been ethically approved via the University of Sheffield's Ethics Review Procedure, as administered by School of Architecture. The University's Research Ethics Committee monitors the application and delivery of the University's Ethics Review Procedure across the University.

13. What if something goes wrong and I wish to complain about the research?

If there are any complaints about the research, please contact the researcher or his supervisor (contact details below). If your complaint has not been handled to your satisfaction, please contact the head of the School of Architecture, Professor Karim Hadjri.

14. Contact for further information

Researcher: Khalid Hamoodh
E: kahamoodh1@sheffield.ac.uk
T:

Supervisor: Prof Steve Fotios
E: steve.fotios@sheffield.ac.uk
T:

Thank you for your time reading this. I hope you be interested and take part in this research project.

Figure E.1. (continued). Participant information sheet for Experiment 1.

Participant Consent Form

Using Lighting to Help Pedestrians Be Safe and Feel Safe

Please tick the appropriate boxes	**Yes**	**No**
Taking Part in the Project		
I have read and understood the project information sheet dated 05/06/2018 or the project has been fully explained to me. (If you will answer No to this question please do not proceed with this consent form until you are fully aware of what your participation in the project will mean.)	☐	☐
I have been given the opportunity to ask questions about the project.	☐	☐
I agree to take part in the project. I understand that taking part in the project will include looking at a series of images on a PC screen and either comparing them in pairs or using a rating scale.	☐	☐
I understand that my taking part is voluntary and that I can withdraw from the study at any time; I do not have to give any reasons for why I no longer want to take part and there will be no adverse consequences if I choose to withdraw.	☐	☐
How my information will be used during and after the project		
I understand my personal details such as name, phone number, address and email address etc. will not be revealed to people outside the project.	☐	☐
I understand and agree that my responses (i.e. results of the experiments) may be quoted in publications, reports, web pages, and other research outputs. I understand that I will not be named in any outputs.	☐	☐
I understand and agree that other authorised researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I understand and agree that other authorised researchers may use my data in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I give permission for the test results that I provide to be deposited with the researcher (khalid Hamoodh) and the project supervisor, Professor Steve Fotios, so that the data may be used for future research and learning.	☐	☐
So that the information you provide can be used legally by the researchers		
I agree to assign the copyright I hold in any materials generated as part of this project to The University of Sheffield.	☐	☐

Name of participant [printed] Signature Date

Name of Researcher [printed] Signature Date

Project contact details for further information:
Principal Investigator: Khalid Hamoodh
E: kahamoodh1@sheffield.ac.uk

Supervisor: Professor Steve Fotios
E: steve.fotios@sheffield.ac.uk

Person who can be contacted in the event of a complaint:
Professor Karim Hadjri
Head of School of Architecture
Western Bank, Sheffield S10 2TN. The Arts Tower, Room 13.07
E: k.hadjri@sheffield.ac.uk
T:

Figure E.2. Consent form for Experiment 1.

UREC Jun 11th 2019

Participant Information Sheet

1. **Research Project Title:** Using Lighting to Help Pedestrians Be Safe and Feel Safe

You are participating in a PhD research project that will be conducted at The University of Sheffield. Before deciding whether to take part it is important you understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Don't hesitate to ask if there is anything that is not clear or if you would like more information – please see contact details provided at the end. Take your time to decide whether or not your wish to participate.

2. **What is the purpose of this project?**

In designing urban streets consideration must be given to how pedestrians' safety and their feelings of safety are continued after dark. The results of this experiment will contribute to ongoing research of current UK lighting standards for residential roads in order to give development recommendations for the British Standards Institution.

3. **Why have I been chosen?**

The study may have reached your email through the volunteer lists for staff or students maintained by CiCS. Other participants will be recruited from within staff and students of the University. The recruitment targets people not aware of the research, and to reach a semi-random sample.

4. **Do I have to take part?**

It is totally up to you to decide whether or not to take part. You can still withdraw at any time without any negative consequences, and you do not have to give a reason.

5. **What will happen to me if I take part? What do I have to do?**

The experiment will be conducted during normal working hours in the lighting laboratory on the 19th floor of the Arts Tower. The whole participation time will not take more than 60 minutes. You will be given a copy of the information sheet and, if appropriate, a signed consent form to keep.

The project does not ask for any personal information except age and gender. Names will be strictly confidential and it will not be asked or mentioned in any phase of the research. All the information needed will be obtained from choosing pictures. In addition, participants will be asked to sign a consent sheet, and they have the right to withdraw at any time of the test without giving any reason and without any harmful consequences.

The experiment requires you to do 3 sessions 10 minutes each, with a 2 minutes break in between. While doing the 3 sessions you will be asked to walk on a treadmill at a very low speed (as you walk normally on the street). The experiment will be started by showing on a

Figure E.3. Participant information sheet for Experiment 2.

screen series of photographs of people in a street scene at night. The different people will all appear to be about 4 metres away from you.

Each session will be used two procedures. First, the photographs will be shown one at a time, each image will be visible for only 0.5 seconds. You will evaluate each image by answering the question "how safe do you feel in the vicinity of this person?" It is a multiple-choice answer on a six-point response scale from "Extremely unsafe" to "Extremely Safe" (you answer by clicking buttons), this procedure will take about 3 minutes. The second procedure will show two images at the same time for unlimited time, and you will choose the safer situation. This procedure asks for 128 comparisons, and it only needs 7 minutes to complete it.

6. What are the possible disadvantages and risks of taking part?

There are no known disadvantages to participation in this experiment. The researcher will be present and thus available to participants at all times if they need any help.

7. What are the possible benefits of taking part?

Participation in this experiment will help us to improve the standards for lighting design, which in turn leads toward a safe and protected community.

8. Will my taking part in this project be kept confidential?

For this research we would like to record your age, gender, and your responses to the test questions. We will use these data in our analyses and will retain the data for future use. Future uses include our further analyses of the data and sharing the results with other researchers. We will ask you to add your name and sign the Consent to Participate form: this is to enable the researcher to confirm that we did seek informed consent to participate. Your name will not be included in our records of results: in other words, all data that you provide will be anonymised and it will not be possible to identify you from the data collected.

All the information that we collect about you during the course of the research will be kept strictly confidential and will only be accessible to members of the research team. You will not be able to be identified in any reports or publications. If you agree to us sharing the information you provide with other researchers (e.g. by making it available in a data archive) then your personal details will not be included.

9. What is the legal basis for processing my personal data?

We will ask to record your age and gender and a visual condition test in order to demonstrate how well the data collected represent the whole population. We will not be

Figure E.3. (continued). Participant information sheet for Experiment 2.

recording your name or contact information. Only collected personal data are gender and age group. Participants' names will be as numbers in the data analysis.

10. What will happen to the data collected, and the results of the research project?

The experimenter will record each participant's gender and age. Each participant will be given a reference number. This means their responses are anonymised which give safety responses.

The data (anonymised test results) will be stored by the experimenter on his own PC with a back-up held by the supervisor.

Due to the nature of this research it is likely that other researchers may find the data collected to be useful in answering future research questions. We will ask for your explicit consent for your data to be shared in this way.

11. Who is organising and funding the research?

The researcher is funded by the Saudi Arabian Cultural Bureau.
The research is organized by The University of Sheffield, School of Architecture.

12. Who is the Data Controller?

The University of Sheffield will act as the Data Controller for this study. This means that the University of Sheffield is responsible for looking after your information and using it properly.

13. Who has ethically reviewed the project?

This project has been ethically approved via the University of Sheffield's Ethics Committee (UREC), as administered by School of Architecture. The University's Research Ethics Committee monitors the application and delivery of the University's Ethics Review Procedure across the University.

14. What if something goes wrong and I wish to complain about the research?

If there are any complaints about the research, please contact the researcher or his supervisor (contact details below). If your complaint has not been handled to your satisfaction, please contact the head of the School of Architecture, Professor Karim Hadjri.

15. Contact for further information

Researcher: Khalid Hamoodh
E: kahamoodh1@sheffield.ac.ukk
T:

Supervisor: Prof Steve Fotios
E: steve.fotios@sheffield.ac.uk
T:

Thank you for your time reading this. I hope you be interested and take part in this research project.

Figure E.3. (continued). Participant information sheet for Experiment 2.

The University Of Sheffield.

Participant Consent Form

[Using Lighting to Help Pedestrians Be Safe and Feel Safe] Consent Form

Please tick the appropriate boxes	Yes	No
Taking Part in the Project		
I have read and understood the project information sheet dated 11/06/2019 or the project has been fully explained to me. (If you will answer No to this question please do not proceed with this consent form until you are fully aware of what your participation in the project will mean.)	☐	☐
I have been given the opportunity to ask questions about the project.	☐	☐
I agree to take part in the project. I understand that taking part in the project will include three test sessions, during which you will be walking on a treadmill while looking at a series of images on a screen.	☐	☐
I understand that my taking part is voluntary and that I can withdraw from the study at any time; I do not have to give any reasons for why I no longer want to take part and there will be no adverse consequences if I choose to withdraw.	☐	☐
How my information will be used during and after the project		
I understand my personal details such as name, phone number, address and email address etc. will not be revealed to people outside the project.	☐	☐
I understand and agree that my words may be quoted in publications, reports, web pages, and other research outputs. I understand that I will not be named in these outputs unless I specifically request this.	☐	☐
I understand and agree that other authorised researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I understand and agree that other authorised researchers may use my data in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I give permission for the test results that I provide to be deposited with the researcher (Khalid Hamoodh) and the project supervisor (Professor Steve Fotios), so it can be used for future research and learning.	☐	☐
So that the information you provide can be used legally by the researchers		
I agree to assign the copyright I hold in any materials generated as part of this project to The University of Sheffield.	☐	☐

Name of participant [printed] Signature Date

Name of Researcher [printed] Signature Date

Project contact details for further information:
Principal Investigator: Khalid Hamoodh
E: kahamoodh1@sheffield.ac.uk
Supervisor: Professor Steve Fotios
E: steve.fotios@sheffield.ac.uk

Person who can be contacted in the event of a complaint:
Professor Karim Hadjri
Head of School of Architecture
Western Bank, Sheffield S10 2TN. The Arts Tower, Room 13.07
E: k.hadjri@sheffield.ac.uk
T:

The template of this consent form has been approved by the University of Sheffield Research Ethics Committee and is available to view here: https://www.sheffield.ac.uk/rs/ethicsandintegrity/ethicspolicy/further-guidance/homepage

Figure E.4. Consent form for Experiment 2.

UREC April 26th 2021

Participant Information Sheet

1. **Research Project Title:** Using Road lighting to Enhance Pedestrians' Safety and their Feelings of Safety.

2. **Invitation paragraph**

You are participating in a PhD research project that will be conducted at The University of Sheffield. Before deciding whether to take part it is important you understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Don't hesitate to ask if there is anything that is not clear or if you would like more information – please see contact details provided at the end. Take your time to decide whether or not your wish to participate.

3. **What is the project's purpose?**

In designing urban streets consideration must be given to how pedestrians' safety and their feelings of safety are continued after dark. The results of this experiment will contribute to ongoing research of current UK lighting standards for residential roads in order to give development recommendations for the British Standards Institution.

4. **Why have I been chosen?**

The study may have reached your email through a previous conversation with the researcher or your participation in previous experiments. Other participants will be recruited from within staff and students of the University. We are aiming for a semi-random sample with representation of society. The only requirements are that you are aged 18 or more and are not considered at excessively vulnerable to Covid-19.

5. **Do I have to take part?**

It is totally up to you to decide whether or not to take part. You can still withdraw at any time without any negative consequences, and you do not have to give a reason. Two reasons prevent you from participating; first, if you have Covid-19 symptoms or you have been in contact with anyone with symptoms in the last 14 days, second, if you need to wear spectacles because the experiment requires wearing eye-tracking glasses.

6. **What will happen to me if I take part? What do I have to do?**

The experiment will be conducted during normal working hours in the lighting laboratory on the 19th floor of the Arts Tower. Participation should not require more than 30 minutes. You will be given a copy of the information sheet and, if appropriate, a signed consent form to keep. Even after signing the consent form, you have the right to withdraw at any time of the test without giving any reason and without any consequences.

Figure E.5. Participant information sheet for Experiment 3.

You will be asked to state your name, age and gender. The age and gender of participants will be used to demonstrate the degree to which the sample represents society. Your name will be recorded only on the consent form, together with an identification number: in all subsequent work your data will be associated only with the ID number, from which it will not be possible to establish your identity. Also, your mobile phone number will be asked to support contact tracing if anyone finds they had Covid's symptoms, it will be confidential only for use this purpose.

During this experiment you will be required to wear an eye-tracking device, which is similar to a normal pair of glasses. The eye-tracking device has inbuilt cameras that capture and record the user's eye movements with the viewable scene. You will be asked to look at a series of images of people and outdoor and indoor views. The photographs will be shown one at a time, for 3 seconds each. In the first part you will be asked only to look at these images: in the second part you will be asked to answer a simple question after each image.

7. What are the possible disadvantages and risks of taking part?

There are no known disadvantages to participation in this experiment. We have taken precautions to safeguard participants and the researcher against risk of Covid-19 infection (see the attached Covid-19 Protection Plan). The researcher will be present at all times and thus available to participants if they need any help.

8. What are the possible benefits of taking part?

Participation in this experiment will help us to improve the standards for lighting design, which in turn leads toward a safe and protected community. As compensation for giving up your time, and to cover any expenses incurred, you will be given £10 at the end of the session.

9. Will my taking part in this project be kept confidential?

For this research we would like to record your age, gender, and your responses to the test questions. We will use these data in our analyses and will retain the data for future use. Future uses include our further analyses of the data and sharing the results with other researchers. We will ask you to add your name and sign the Consent to Participate form: this is to enable the researcher to confirm that we did seek informed consent to participate. Your name will not be included in our records of results: in other words, all data that you provide will be anonymised and it will not be possible to identify you from the data collected.

All the information that we collect about you during the course of the research will be kept strictly confidential and will only be accessible to members of the research team. You will not be able to be identified in any reports or publications. If you agree to us sharing the information you provide with other researchers (e.g. by making it available in a data archive) then your personal details will not be included.

Figure E.5. (continued). Participant information sheet for Experiment 3.

UREC April 26th 2021

10. What is the legal basis for processing my personal data?

We will ask to record your age and gender and a visual condition test in order to demonstrate how well the data collected represent the whole population. We will not be recording your name or contact information. Only collected personal data are gender and age group. Participants' names will be as numbers in the data analysis.

11. What will happen to the data collected, and the results of the research project?

The experimenter will record each participant's gender and age. Each participant will be given a reference number. This means their responses are anonymised which give safety responses. The data (anonymised test results) will be stored by the experimenter on his own PC with a back-up held by the supervisor. Due to the nature of this research it is likely that other researchers may find the data collected to be useful in answering future research questions. We will ask for your explicit consent for your data to be shared in this way.

12. Who is organising and funding the research?

The researcher is funded by the Saudi Arabian Cultural Bureau, and organized by a lighting group member at The University of Sheffield, School of Architecture.

13. Who is the Data Controller?

The University of Sheffield will act as the Data Controller for this study. This means that the University of Sheffield is responsible for looking after your information and using it properly.

14. Who has ethically reviewed the project?

This project has been ethically approved via the University of Sheffield's Ethics Committee (UREC), as administered by School of Architecture. The University's Research Ethics Committee monitors the application and delivery of the University's Ethics Review Procedure across the University.

15. What if something goes wrong and I wish to complain about the research?

If there are any complaints about the research, please contact the researcher or his supervisor (contact details below). If your complaint has not been handled to your satisfaction, please contact the head of the School of Architecture, Professor Karim Hadjri.

16. Contact for further information

Researcher: Khalid Hamoodh
E: kahamoodh1@sheffield.ac.uk
T:

Supervisor: Prof Steve Fotios
E: steve.fotios@sheffield.ac.uk
T:

Thank you for your time reading this. I hope you be interested and take part in this research project.

Figure E.5. (continued). Participant information sheet for Experiment 3.

The University Of Sheffield.

Participant Consent Form

[Using Road lighting to Enhance Pedestrians' Safety and their Feelings of Safety]
Consent Form

Please tick the appropriate boxes	**Yes**	**No**
Taking Part in the Project		
I have read and understood the project information sheet dated 01/04/2021 or the project has been fully explained to me. (If you will answer No to this question please do not proceed with this consent form until you are fully aware of what your participation in the project will mean.)	☐	☐
I have been given the opportunity to ask questions about the project.	☐	☐
I agree to take part in the project. I understand that taking part in the project will include two phases, during which you will be wearing eye-tracking glasses, while looking at a series of images on a screen.	☐	☐
I understand that my taking part is voluntary and that I can withdraw from the study at any time; I do not have to give any reasons for why I no longer want to take part and there will be no adverse consequences if I choose to withdraw.	☐	☐
How my information will be used during and after the project		
I understand my personal details such as name, phone number, address and email address etc. will not be revealed to people outside the project.	☐	☐
I understand and agree that my words may be quoted in publications, reports, web pages, and other research outputs. I understand that I will not be named in these outputs unless I specifically request this.	☐	☐
I understand and agree that other authorised researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I understand and agree that other authorised researchers may use my data in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I give permission for the test results that I provide to be deposited with the researcher (Khalid Hamoodh) and the project supervisor (Professor Steve Fotios), so it can be used for future research and learning.	☐	☐
So that the information you provide can be used legally by the researchers		
I agree to assign the copyright I hold in any materials generated as part of this project to The University of Sheffield.	☐	☐

Name of participant [printed] Signature Date

Name of Researcher [printed] Signature Date

Project contact details for further information:
Principal Investigator: Khalid Hamoodh
E: kahamoodh1@sheffield.ac.uk
Supervisor: Professor Steve Fotios
E: steve.fotios@sheffield.ac.uk

Person who can be contacted in the event of a complaint:
Professor Karim Hadjri
Head of School of Architecture
Western Bank, Sheffield S10 2TN. The Arts Tower, Room 13.07
E: k.hadjri@sheffield.ac.uk
T:

The template of this consent form has been approved by the University of Sheffield Research Ethics Committee and is available to view here: https://www.sheffield.ac.uk/rs/ethicsandintegrity/ethicspolicy/further-guidance/homepage

Figure E.6. Consent form for Experiment 3.

Preventative Plan from Covid-19

Research: Using Road lighting to Enhance Pedestrians' Safety and their Feelings of Safety.
Implementing: year 2021

We understand the COVID-19 situation and the participants' harm in this experiment where they might get infected with the virus while a participant. All the government guidelines and regulations as well the NHS information and advice about coronavirus (COVID-19) are remains in place to work safely in this experiment. This experiment will be conducted by an individual participant following these precautions steps:

Few days before arriving (how to prevent contagion):

1. This document will be included with the information sheet sent to the participants before they arrive.
2. It will mention in the recruiting method and in the information sheet; any participant having Covid-19 symptoms or who has been in contact with anyone with symptoms in last 14 days will not be invited to participate.
3. The experimenter will be tested frequently (every three days) at the Covid-19 Testing Centre (https://www.sheffield.ac.uk/coronavirus/safety/our-covid-19-testing-centre). He will then be able to confirm to participants that he is clear and doesn't have any symptoms.

When they arrive (how to get in safely):

1. The experimenter will be waiting for participants on the 19th floor of the Arts Tower, the location of the laboratory. The experiment will be set up ready for the participant and surfaces sanitised.
2. The laboratory is on the 19^{th} floor. To reach this floor, participants should use only the right-hand lift. it will be reminded in the reminder email a day before their time slot.
3. Upon the participant arriving to the lab, the experimenter will use the hands sanitizer in front of the participant (just to illustrate the way) and the participant will repeat that.
4. Participants' mobile phone number will be asked to support contact tracing if anyone finds they had Covid's symptoms. It will be confidential only for use this purpose.

While doing the experiment (how to work safely):

1. Test participants will wear a mask for the whole period, from entry to departure of the Arts Tower. They will be asked to sanitise their hands at the beginning and at the end of the test session. The laboratory holds a supply of masks and sanitiser.

Figure E.7. Prevention plan from Covid-19 form used in Experiment 3.

2. The eye-tracking glasses will be sanitized in front of the participant, so they will be happy and feel safer about participation.
3. The participant will wear the glasses by them self via a demonstration video display on the screen, so the experimenter will never need close contact with the participant.
4. Participants are not required to touch anything other than the eye tracking glasses; their responses will be oral. In case if they touch anything the hands sanitizer and antibacterial wipes are available in the lab to use.

When they finish the experiment (how to get out safely):

1. Participant will sanitize their hands before they leave.
2. The experimenter will remind the participant how to get out of the building (same way of getting in).

Thank you for your interests to be a part of this study. We did our best to come safe, work safe, and leave safe. Wishing you happy participation.

Researcher: Khalid Hamoodh
E: kahamoodh1@sheffield.ac.ukk
T:

Supervisor: Prof Steve Fotios
E: steve.fotios@sheffield.ac.uk
T:

Figure E.7. (continued). Prevention plan from Covid-19 form used in Experiment 3.

Appendix F. Photo release form

Photo Release Form

I understand that the Khalid Hamoodh is a PhD student at Sheffield University studying how we evaluate other people when walking after dark. The research requires to take photos of people in different poses. The photos will only be used for the research purposes such as illustration, analysis, thesis, and journals publication.

Please check the correct box and complete the information below to whether you do wish or do not wish to grant Khalid Hamoodh to use your taken photos.

☐ I DO grant permission for my taken photo.

☐ I DO NOT grant permission for the use of my taken photo.

Name ____

Signature ______________________________ Date ______________

Figure F.1. Example of photo release form that was signed by the actors of Experiments 1 and 2.

Appendix G. Experiments raw data

Table G.1. Responses of Experiment 1 for each test image.

Participant ID	Test image no.															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1	2	2	2	5	3	2	3	5	4	4	3	5	3	2	2	5
2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
3	4	4	5	3	5	4	5	4	4	3	4	4	4	3	3	3
4	5	5	5	5	4	4	5	3	3	4	5	5	5	4	4	3
5	3	3	4	4	4	3	3	2	3	3	3	3	4	3	4	4
6	3	2	4	3	3	2	3	2	5	2	4	5	4	3	5	3
7	4	4	4	4	3	3	3	4	3	4	3	3	3	3	3	3
8	3	2	5	4	2	1	3	3	4	3	5	3	3	2	5	3
9	3	3	4	4	2	2	3	3	4	4	4	3	3	2	4	3
10	3	2	4	4	3	2	4	4	4	4	4	2	4	3	5	5
11	2	2	4	4	2	3	4	3	3	4	5	4	2	2	5	3
12	5	4	4	3	4	3	4	3	4	3	3	3	3	3	4	4
13	5	4	5	3	5	4	5	3	5	5	5	4	5	5	5	5
14	5	4	5	5	5	5	5	5	4	4	4	4	4	5	5	4
15	5	4	5	4	5	3	4	4	4	4	5	4	5	3	4	4
16	4	3	5	4	5	3	4	4	5	2	4	2	4	4	4	3
17	5	2	5	4	3	2	4	2	5	3	5	2	5	4	5	4
18	4	2	5	4	1	2	4	4	4	2	4	2	2	2	5	5
19	4	4	4	4	3	3	4	3	4	4	4	3	4	4	4	4
20	5	5	4	5	4	4	4	4	5	5	5	4	4	4	4	4
21	4	4	5	4	4	4	5	5	4	2	5	2	5	4	5	3
22	3	2	4	4	3	3	4	2	4	3	4	3	4	3	5	5
23	3	2	4	2	2	4	4	2	3	2	4	2	3	3	4	4
24	5	4	5	4	4	4	4	4	4	4	4	3	5	4	3	4
25	3	2	5	4	1	3	3	2	5	1	4	1	3	3	5	5
26	4	4	4	4	3	4	4	4	5	3	5	3	4	3	5	5
27	2	4	3	3	2	3	2	2	3	3	4	3	4	2	4	4
28	4	3	4	5	3	3	3	3	4	4	5	3	5	3	5	4
29	3	3	5	5	2	3	5	4	5	2	5	5	2	2	5	5
30	4	5	5	5	5	5	5	5	5	3	5	3	2	2	5	4
31	3	3	3	3	3	2	3	3	4	2	4	2	3	4	3	3
32	4	4	4	2	4	2	4	2	4	4	4	4	4	4	4	4

Table G.2. Responses of Experiment 2 for each test image in face block.

Scene	Participant ID	Test image no.															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4	5	4	3	1	6	4	2	3	5	5	5	2	5	2	4	2
	9	5	3	4	2	5	4	4	4	5	5	5	3	5	4	4	3
	10	5	2	2	1	5	4	2	1	6	3	5	2	6	5	4	2
	11	5	4	3	3	4	3	3	3	6	5	5	3	6	5	4	4
	12	4	2	3	2	4	3	3	2	5	4	4	2	5	3	4	2
	14	5	4	5	4	5	4	4	3	5	5	5	5	6	5	5	4
	15	4	3	4	2	4	3	4	2	4	3	4	2	4	3	4	2
	16	5	3	4	2	4	4	2	2	5	3	5	3	5	4	5	2
	19	5	4	2	3	5	4	2	2	6	5	2	3	6	5	3	3
	21	4	4	1	1	4	4	2	1	5	5	4	3	5	4	3	2
	23	4	3	3	1	6	5	1	1	6	5	3	1	6	5	3	1
	24	6	2	2	1	6	4	3	2	6	3	2	2	6	4	3	1
	26	4	4	6	6	6	5	4	6	6	5	6	6	6	6	6	5
	28	4	3	3	3	4	4	3	3	5	4	4	3	5	4	4	3
	29	5	3	2	1	5	1	2	1	6	6	3	3	5	2	3	1
	32	4	2	2	1	4	1	2	2	4	3	4	3	3	2	3	1
	33	6	3	3	2	6	3	4	2	6	5	4	4	6	4	5	4
	36	4	4	3	1	4	3	3	2	5	4	3	2	5	4	3	2
	39	4	3	4	3	4	2	3	2	5	5	5	3	3	3	3	2
	40	6	4	4	1	6	2	3	1	6	3	5	3	6	2	3	2
	42	5	2	1	1	5	1	1	1	6	4	3	2	5	4	2	2
	44	5	3	3	1	6	4	2	1	5	3	2	1	4	2	2	1
2	1	5	4	3	2	4	3	3	3	5	5	3	2	5	4	3	3
	2	5	4	5	3	5	4	5	4	5	5	5	4	5	5	5	4
	3	6	3	4	1	6	3	5	1	6	4	4	2	6	4	5	2
	5	4	3	2	1	3	3	2	1	5	4	3	3	4	3	2	2
	6	5	2	2	1	6	3	3	1	6	1	4	1	6	3	4	1
	7	5	2	1	1	5	2	1	1	5	2	3	1	5	3	2	1
	8	4	3	3	3	5	4	3	2	4	3	3	3	4	3	2	3
	13	6	2	4	2	6	3	5	1	6	2	3	2	6	2	4	2
	17	3	3	4	1	3	2	2	1	5	4	4	4	3	4	3	3
	18	4	3	4	1	4	3	3	1	5	3	3	2	5	5	2	2
	20	5	4	4	4	5	5	3	4	5	4	4	5	5	4	4	3
	22	5	4	3	3	6	3	3	3	5	5	4	2	5	4	4	3
	25	3	1	2	1	4	2	2	1	6	5	5	4	6	4	4	3
	27	5	2	4	1	5	2	2	2	6	3	2	1	5	3	3	2
	30	4	4	3	3	6	4	4	3	5	5	5	5	6	5	5	5
	31	5	3	4	2	5	3	4	1	5	3	4	2	5	3	4	1
	34	4	3	3	2	3	3	2	2	4	4	4	3	5	3	4	2
	35	5	3	4	1	5	4	4	2	6	4	3	1	5	3	3	2
	37	3	2	2	1	4	3	2	2	4	2	2	1	4	3	2	1
	38	5	5	3	3	5	4	5	4	5	5	5	5	5	6	5	5
	41	6	3	3	3	6	3	3	1	6	5	4	1	6	3	4	1
	43	3	2	2	1	4	2	1	1	3	2	2	2	4	2	2	1

Table G.3. Responses of Experiment 2 for each test image in hands block.

Scene	Participant ID	Test image no.															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4	6	6	6	5	6	6	6	6	6	6	5	6	6	5	6	6
	9	5	5	5	4	5	5	4	3	6	5	5	5	6	5	5	4
	10	5	5	5	3	5	4	3	3	6	6	5	5	5	5	6	5
	11	5	5	4	3	4	5	5	4	5	5	5	4	5	5	5	6
	12	4	3	4	4	4	4	4	4	6	5	5	5	4	4	4	4
	14	5	5	5	5	5	5	5	5	5	6	5	5	6	6	6	6
	15	4	4	3	2	4	4	3	2	4	5	3	2	5	5	3	2
	16	2	3	2	5	4	4	4	4	3	4	2	5	4	4	3	3
	19	6	4	4	5	6	5	6	5	6	6	6	5	6	5	6	5
	21	5	3	5	5	4	3	5	3	5	3	5	5	5	3	5	5
	23	5	4	5	4	5	5	5	3	5	6	5	6	6	6	6	3
	24	6	5	4	4	3	5	4	3	6	4	4	4	6	5	4	4
	26	3	3	3	3	5	5	5	5	6	6	5	5	6	6	5	4
	28	3	3	3	3	4	5	4	5	5	5	5	5	5	5	5	5
	29	4	6	6	6	6	5	5	4	6	6	4	6	4	6	4	5
	32	3	3	3	3	3	3	3	3	3	2	4	4	2	2	2	2
	33	5	6	6	3	3	6	4	3	6	6	6	4	6	6	4	5
	36	4	3	3	3	4	3	5	5	5	5	4	5	3	4	3	4
	39	3	4	3	4	4	4	4	4	5	5	4	5	4	4	4	4
	40	6	5	4	3	6	6	4	3	6	6	5	4	6	6	5	3
	42	5	3	4	3	3	3	5	4	6	6	6	5	5	4	5	4
	44	5	4	5	5	4	2	2	1	6	5	6	2	4	4	2	3
2	1	4	3	4	2	2	4	3	4	3	4	2	4	4	5	5	5
	2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	3	6	5	5	2	6	4	4	2	6	4	5	2	6	6	5	3
	5	4	3	3	3	4	3	4	3	5	5	5	5	5	5	4	4
	6	4	5	4	3	5	3	4	3	6	4	4	4	4	4	4	4
	7	5	3	3	1	5	2	3	2	5	3	4	2	5	3	4	2
	8	4	4	4	4	5	5	4	4	5	4	4	5	5	4	4	4
	13	6	6	6	6	6	6	5	5	6	6	6	5	6	6	6	6
	17	3	4	3	3	3	4	2	3	5	5	4	4	4	5	3	4
	18	5	4	4	4	6	5	5	5	6	5	5	4	6	5	5	5
	20	5	5	5	4	5	4	5	4	5	4	5	5	5	4	5	4
	22	4	4	4	3	5	4	5	4	5	5	4	4	5	5	5	4
	25	3	2	3	1	5	3	4	4	6	5	6	4	6	6	5	4
	27	5	3	3	3	4	3	4	4	6	4	4	3	4	4	6	2
	30	5	5	6	5	5	6	6	5	6	6	6	6	6	6	6	6
	31	5	5	5	3	5	5	5	2	5	5	6	3	5	5	5	3
	34	4	4	4	3	3	4	4	3	3	5	5	4	4	5	3	4
	35	5	4	3	4	5	4	5	3	5	4	5	2	5	5	5	2
	37	3	3	2	1	3	4	3	1	5	5	4	2	4	5	4	2
	38	4	5	5	4	4	5	4	4	5	6	5	6	6	5	5	5
	41	5	6	6	3	5	6	6	4	6	6	6	5	6	6	4	4
	43	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5

Table G.4. Responses of Experiment 2 for each test image in mixed block.

Scene	Participant ID	Test image no.															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4	5	2	5	1	5	1	5	1	6	4	6	1	5	2	5	3
	9	6	1	4	2	5	1	4	2	6	2	5	4	6	3	3	3
	10	5	1	5	1	5	1	5	1	6	2	6	1	6	1	6	1
	11	4	3	4	3	5	3	4	3	6	4	6	4	6	3	6	3
	12	4	2	3	2	5	2	3	2	5	3	4	2	5	3	4	2
	14	5	5	5	5	6	5	6	5	6	5	6	5	6	5	6	5
	15	3	2	2	2	4	3	3	2	4	3	3	2	4	2	3	2
	16	4	1	4	2	6	2	5	2	4	2	4	2	4	1	6	1
	19	5	1	4	1	5	2	6	1	6	1	6	2	6	1	6	1
	21	5	2	5	2	5	2	5	2	6	4	5	3	5	3	5	2
	23	4	1	2	1	5	1	2	1	6	1	3	1	6	2	3	1
	24	6	2	4	1	6	2	5	1	6	2	4	1	6	2	4	1
	26	3	2	3	1	6	3	3	2	6	4	3	2	6	4	3	2
	28	3	2	3	3	4	2	4	2	5	3	5	3	5	3	5	3
	29	6	1	6	1	6	3	6	1	6	1	6	2	6	1	6	1
	32	4	2	4	2	3	2	4	2	5	4	5	4	3	4	4	4
	33	5	2	4	1	5	2	3	2	6	3	5	2	6	3	4	2
	36	3	3	3	2	5	2	5	2	6	3	5	3	4	3	4	3
	39	4	2	4	2	4	2	4	2	5	3	5	3	4	2	4	2
	40	6	2	4	1	6	2	4	1	6	2	4	1	6	2	5	1
	42	5	1	3	1	5	1	3	1	6	3	6	3	5	3	3	2
	44	5	2	3	1	2	2	3	1	5	2	4	1	5	2	3	1
2	1	5	2	3	2	4	2	3	3	6	3	6	1	5	3	3	2
	2	5	2	5	2	6	2	5	2	5	5	5	5	5	4	5	3
	3	6	2	4	1	6	2	3	1	6	2	4	1	6	2	2	1
	5	4	2	3	2	3	2	3	2	5	2	5	2	5	2	5	2
	6	4	2	4	1	5	2	3	1	6	2	4	1	5	2	3	1
	7	6	1	2	1	5	1	4	1	6	1	3	1	5	1	3	1
	8	3	3	4	2	5	2	5	3	5	3	5	3	4	3	4	2
	13	5	2	5	1	6	3	5	2	6	2	3	1	6	2	5	2
	17	3	2	3	2	3	2	3	2	6	4	5	3	4	3	5	3
	18	6	1	5	2	6	1	5	2	6	3	5	3	6	3	5	3
	20	5	2	3	3	5	3	5	3	5	2	4	2	5	4	4	4
	22	4	3	4	2	5	3	4	2	5	4	4	3	5	3	4	3
	25	3	2	3	1	2	2	4	2	6	4	6	4	6	4	6	4
	27	4	1	2	1	4	1	2	3	5	1	2	1	5	1	3	1
	30	3	3	4	3	5	3	5	3	5	5	5	5	5	5	5	5
	31	5	1	5	1	5	1	5	1	5	2	5	1	5	1	5	1
	34	5	2	4	2	4	3	3	2	5	3	5	2	5	3	4	3
	35	5	1	2	1	5	3	3	1	6	2	5	1	5	2	2	1
	37	2	1	1	1	3	1	2	1	3	1	2	1	4	1	3	1
	38	4	2	5	3	4	2	4	2	6	4	6	4	6	3	6	4
	41	5	4	5	3	6	4	5	2	6	3	6	3	5	4	4	3
	43	5	1	4	2	5	2	4	2	6	2	6	2	5	2	5	2

Appendix H. Normality check

Table H.1. Normality profile for Experiment 1.

		Image 1	Image 2	Image 3	Image 4
Central Tendency	Mean	3.75	3.28	4.31	3.91
	95% CI of mean*	3.41-4.09	2.91-3.65	4.05-4.58	3.61-4.20
	Median	4.00	3.50	4.00	4.00
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	Yes
Graphical	Histogram	Near	No	No	Yes
	Box Plot	Near	No	No	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness**	-0.180	-0.036	-1.096	-0.576
	(within ±0.5)	Yes	Yes	No	No
	Kurtosis**	-0.872	-1.292	1.681	0.234
	(within ±1.0)	Yes	No	No	Yes
NORMALITY?	(Yes if both are yes)	Yes	No	No	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	0.003	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	0.002	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.1. (continued). Normality profile for Experiment 1.

		Image 5	**Image** 6	**Image** 7	**Image** 8
Central Tendency	Mean	3.31	3.09	3.88	3.34
	95% CI of mean*	2.89-3.74	2.75-3.44	3.59-4.16	2.98-3.71
	Median	3.00	3.00	4.00	3.00
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	Yes
Graphical	Histogram	Near	Yes	Near	No
	Box Plot	Near	Yes	No	Near
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	Yes	No	No
Measures of dispersion	Skewness**	-0.152	0.035	-0.181	0.051
	(within ±0.5)	Yes	Yes	No	Yes
	Kurtosis**	-0.744	-0.433	-0.459	-1.061
	(within ±1.0)	Yes	Yes	Yes	No
NORMALITY?	(Yes if both are yes)	Yes	Yes	No	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	0.023	0.003	<0.001	0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	0.015	0.010	0.001	0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**Yes**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.1. (continued). Normality profile for Experiment 1.

		Image 9	**Image** 10	**Image** 11	**Image** 12
Central Tendency	Mean	4.09	3.25	4.25	3.22
	95% CI of mean*	3.85-4.34	2.90-3.60	4.01-4.49	2.84-3.59
	Median	4.00	3.00	4.00	3.00
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	Yes
Graphical	Histogram	Near	Near	Near	Yes
	Box Plot	No	Near	No	Near
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness**	-0.123	-0.325	-0.340	0.083
	(within ±0.5)	Yes	Yes	Yes	Yes
	Kurtosis**	-0.768	-0.570	-0.698	-0.527
	(within ±1.0)	Yes	Yes	Yes	Yes
NORMALITY?	(Yes if both are yes)	Yes	Yes	Yes	Yes
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	<0.001	0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	0.003	<0.001	0.013
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.1. (continued). Normality profile for Experiment 1.

		Image 13	**Image** 14	**Image** 15	**Image** 16
Central Tendency	Mean	3.72	3.19	4.28	3.94
	95% CI of mean*	3.37-4.06	2.86-3.51	3.99-4.57	3.66-4.21
	Median	4.00	3.00	4.00	4.00
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	Yes
Graphical	Histogram	Near	No	No	No
	Box Plot	No	No	No	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness**	-0.323	0.182	-0.961	0.107
	(within ±0.5)	Yes	Yes	No	Yes
	Kurtosis**	-0.714	-0.768	0.453	-1.203
	(within ±1.0)	Yes	Yes	Yes	No
NORMALITY?	(Yes if both are yes)	Yes	Yes	No	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	0.001	0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval
** Z–score of skewness (and kurtosis) were provided by SPSS

Table H.2. Normality profile for Experiment 2 (face block).

		Image 1	Image 2	Image 3	Image 4
Central Tendency	Mean	4.64	3.05	3.09	1.91
	95% CI of mean*	4.4-4.9	2.8-3.3	2.8-3.4	1.6-2.3
	Median	5	3	3	2
NORMALITY?	(Yes if median is in 95% CI for mean)	No	Yes	Yes	Yes
Graphical	Histogram	Yes	Yes	Yes	No
	Box Plot	No	Yes	Near	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	Yes	Yes	No
Measures of dispersion	Skewness**	-0.205	-0.090	0.146	1.377
	(within ±0.5)	Yes	Yes	Yes	No
	Kurtosis**	-0.395	-0.500	0.090	2.329
	(within ±1.0)	Yes	Yes	Yes	No
NORMALITY?	(Yes if both are yes)	Yes	Yes	Yes	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	0.002	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	<0.001	0.006	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**Yes**	**Yes**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.2. (continued). Normality profile for Experiment 2 (face block).

		Image 5	**Image** 6	**Image** 7	**Image** 8
Central Tendency	Mean	4.84	3.18	2.86	2.02
	95% CI of mean*	4.6-5.1	2.9-3.5	2.5-3.2	1.7-2.4
	Median	5	3	3	2
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	Yes
Graphical	Histogram	No	Near	Near	No
	Box Plot	No	No	Yes	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	Yes	No
Measures of dispersion	Skewness**	-0.199	-0.383	0.281	1.296
	(within ±0.5)	Yes	Yes	Yes	No
	Kurtosis**	-0.985	-0.202	-0.552	1.918
	(within ±1.0)	Yes	Yes	Yes	No
NORMALITY?	(Yes if both are yes)	Yes	Yes	Yes	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	=0.001	0.002	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**Yes**	**No**

* CI = Confidence Interval
** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.2. (continued). Normality profile for Experiment 2 (face block).

		Image 9	**Image** 10	**Image** 11	**Image** 12
Central Tendency	Mean 95% CI of mean* Median	5.23 5.0-5.5 5	3.93 3.6-4.3 4	3.80 3.5-4.1 4	2.66 2.3-3.1 3
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	Yes
Graphical	Histogram Box Plot	No No	Near Near	Near Yes	Near No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	Yes	No
Measures of dispersion	Skewness** (within ±0.5) Kurtosis** (within ±1.0)	-0.755 No 0.473 Yes	-0.536 No -0.488 Yes	-0.169 Yes -0.855 Yes	0.685 No 0.029 Yes
NORMALITY?	(Yes if both are yes)	No	No	Yes	No
Statistical tests					
Kolmogorov-Smirnov	level of significance (not normal if p<0.05)	<0.001 No	<0.001 No	<0.001 No	<0.001 No
Shapiro-Wilks	level of significance (not normal if p<0.05)	<0.001 No	<0.001 No	=0.001 No	=0.001 No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**Yes**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.2. (continued). Normality profile for Experiment 2 (face block).

		Image 13	**Image** 14	**Image** 15	**Image** 16
Central Tendency	Mean	5.07	3.66	3.55	2.34
	95% CI of mean*	4.8-5.3	3.3-4.0	3.2-3.9	2.0-2.7
	Median	5	4	4	2
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	No	Yes
Graphical	Histogram	No	Near	Near	No
	Box Plot	No	No	No	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness**	-0.795	0.180	0.130	0.781
	(within ±0.5)	No	Yes	Yes	No
	Kurtosis**	0.181	-0.621	-0.636	-0.008
	(within ±1.0)	Yes	Yes	Yes	Yes
NORMALITY?	(Yes if both are yes)	No	Yes	Yes	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	=0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	0.003	0.002	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

Table H.3. Normality profile for Experiment 2 (hands block).

		Image 1	Image 2	Image 3	Image 4
Central Tendency	Mean	4.50	4.25	4.16	3.57
	95% CI of mean*	4.2-4.8	3.9-4.6	3.8-4.5	3.2-4.0
	Median	5	4	4	3
NORMALITY?	(Yes if median is in 95% CI for mean)	No	Yes	Yes	Yes
Graphical	Histogram	Yes	Near	Near	Near
	Box Plot	No	Yes	Yes	No
NORMALITY?	(Yes if both are yes ,2 nears=No)	No	Yes	Yes	No
Measures of dispersion	Skewness**	-0.439	0.020	0.089	-0.125
	(within ±0.5)	Yes	Yes	Yes	Yes
	Kurtosis**	-0.367	-1.082	-0.845	-0.177
	(within ±1.0)	Yes	No	Yes	Yes
NORMALITY?	(Yes if both are yes)	Yes	No	Yes	Yes
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	0.003	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	=0.001	0.002	0.005
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**Yes**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.3. (continued). Normality profile for Experiment 2 (hands block).

		Image 5	**Image** 6	**Image** 7	**Image** 8
Central Tendency	Mean 95% CI of mean* Median	4.50 4.2-4.8 5	4.34 4.0-4.7 4	4.39 4.1-4.7 4	3.66 3.3-4.0 4
NORMALITY?	(Yes if median is in 95% CI for mean)	No	Yes	Yes	Yes
Graphical	Histogram Box Plot	Near No	Near No	Yes No	Yes No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness** (within ±0.5) Kurtosis** (within ±1.0)	-0.341 Yes -0.493 Yes	-0.268 Yes -0.551 Yes	-0.305 Yes -0.074 Yes	-0.200 Yes -0.248 Yes
NORMALITY?	(Yes if both are yes)	Yes	Yes	Yes	Yes
Statistical tests Kolmogorov-Smirnov Shapiro-Wilks	 level of significance (not normal if p<0.05) level of significance (not normal if p<0.05)	 <0.001 No =0.001 No	 <0.001 No 0.002 No	 <0.001 No =0.001 No	 <0.001 No 0.007 No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.3. (continued). Normality profile for Experiment 2 (hands block).

		Image 9	**Image** 10	**Image** 11	**Image** 12
Central Tendency	Mean	5.16	4.89	4.70	4.36
	95% CI of mean*	4.9-5.5	4.6-5.2	4.4-5.0	4.0-4.7
	Median	5	5	5	5
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	No
Graphical	Histogram	No	No	Near	Near
	Box Plot	No	No	No	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness**	-1.468	-0.987	-0.873	-0.773
	(within ±0.5)	No	No	No	No
	Kurtosis**	1.857	0.810	1.336	-0.001
	(within ±1.0)	No	Yes	No	Yes
NORMALITY?	(Yes if both are yes)	No	No	No	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.3. (continued). Normality profile for Experiment 2 (hands block).

		Image 13	**Image** 14	**Image** 15	**Image** 16
Central Tendency	Mean	5.02	4.91	4.61	4.02
	95% CI of mean*	4.7-5.3	4.6-5.1	4.3-4.9	3.6-4.4
	Median	5	5	5	4
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	No	Yes
Graphical	Histogram	No	Near	Near	Yes
	Box Plot	No	No	No	Near
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	Yes
Measures of dispersion	Skewness**	-0.894	-0.884	-0.479	-0.404
	(within ±0.5)	No	No	Yes	Yes
	Kurtosis**	0.871	1.014	-0.140	-0.355
	(within ±1.0)	Yes	No	Yes	Yes
NORMALITY?	(Yes if both are yes)	No	No	Yes	Yes
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	<0.001	=0.001	0.004
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**Yes**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

Table H.4. Normality profile for Experiment 2 (mixed block).

		Image 1	**Image 2**	**Image 3**	**Image 4**
Central Tendency	Mean	4.48	1.91	3.73	1.76
	95% CI of mean*	4.2-4.8	1.7-2.17	3.4-4.1	1.5-2.0
	Median	5	2	4	2
NORMALITY?	(Yes if median is in 95% CI for mean)	No	Yes	Yes	Yes
Graphical	Histogram	Near	No	Yes	No
	Box Plot	No	No	Near	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	Yes	No
Measures of dispersion	Skewness**	-0.321	1.338	-0.335	1.421
	(within ±0.5)	Yes	No	Yes	No
	Kurtosis**	-0.639	3.054	-0.226	3.083
	(within ±1.0)	Yes	No	Yes	No
NORMALITY?	(Yes if both are yes)	Yes	No	Yes	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	=0.001	<0.001	0.005	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**Yes**	**No**

* CI = Confidence Interval
** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.4. (continued). Normality profile for Experiment 2 (mixed block).

		Image 5	**Image** 6	**Image** 7	**Image** 8
Central Tendency	Mean	4.77	2.11	4.02	1.86
	95% CI of mean*	4.5-5.1	1.9-2.4	3.7-4.4	1.6-2.1
	Median	5	2	4	2
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	Yes
Graphical	Histogram	Near	No	Near	No
	Box Plot	Near	No	Yes	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	Yes	No
Measures of dispersion	Skewness**	-0.932	0.888	-0.047	1.310
	(within ±0.5)	No	No	Yes	No
	Kurtosis**	0.482	1.650	-0.880	3.463
	(within ±1.0)	Yes	No	Yes	No
NORMALITY?	(Yes if both are yes)	No	No	Yes	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	<0.001	=0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**Yes**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.4. (continued). Normality profile for Experiment 2 (mixed block).

		Image 9	**Image** 10	**Image** 11	**Image** 12
Central Tendency	Mean	5.52	2.75	4.70	2.32
	95% CI of mean*	5.3-5.7	2.4-3.1	4.4-5.1	2.0-2.7
	Median	6	3	5	2
NORMALITY?	(Yes if median is in 95% CI for mean)	No	Yes	Yes	Yes
Graphical	Histogram	No	Near	Near	No
	Box Plot	No	Yes	No	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	Yes	No	No
Measures of dispersion	Skewness**	-1.586	0.226	-0.687	0.620
	(within ±0.5)	No	Yes	No	No
	Kurtosis**	2.813	-0.727	-0.209	-0.591
	(within ±1.0)	No	Yes	Yes	Yes
NORMALITY?	(Yes if both are yes)	No	Yes	No	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	<0.001	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	0.003	<0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**Yes**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.4. (continued). Normality profile for Experiment 2 (mixed block).

		Image 13	**Image** 14	**Image** 15	**Image** 16
Central Tendency	Mean	5.16	2.57	4.30	2.23
	95% CI of mean*	4.9-5.4	2.2-3.0	3.9-4.7	1.9-2.6
	Median	5	3	4	2
NORMALITY?	(Yes if median is in 95% CI for mean)	Yes	Yes	Yes	Yes
Graphical	Histogram	No	Near	No	No
	Box Plot	No	No	Yes	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness**	-0.602	0.272	-0.069	0.660
	(within ±0.5)	No	Yes	Yes	No
	Kurtosis**	-0.094	-0.435	-1.002	-0.318
	(within ±1.0)	Yes	Yes	No	Yes
NORMALITY?	(Yes if both are yes)	No	Yes	No	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.001	=0.001	=0.001	<0.001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.001	0.002	0.002	<0.001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

Table H.5. Normality profile for Experiment 3.

		Primary	Left position	Right Position	Dog
Central Tendency	Mean	350.51	315.87	308.84	308.24
	95% CI of mean*	300.47 - 400.55	274.29 - 357.45	269.25 - 348.43	267.99 - 348.49
	Median	233.2	216.6	199.90	216.30
NORMALITY?	(Yes if median is in 95% CI for mean)	No	No	No	No
Graphical	Histogram	No	No	No	No
	Box Plot	No	No	No	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness**	3.860	3.810	3.726	4.721
	(within ±0.5)	No	No	No	No
	Kurtosis**	19.292	19.275	19.198	30.99
	(within ±1.0)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.0001	<0.0001	<0.0001	<0.0001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.0001	<0.0001	<0.0001	<0.0001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval
** Z–score of skewness (and kurtosis) were provided by SPSS

TABLE CONTINUES ON THE NEXT PAGE

Table H.5. (continued). Normality profile for Experiment 3.

		Pair of people	Vehicles	Obscured face	Obscured body
Central Tendency	Mean*	267.48	309.97	301.67	411.49
	95% CI of mean	241.94 - 293.02	267.20 - 352.7358	273.21 - 330.14	358.23 - 464.74
	Median	199.90	216.40	233.00	266.30
NORMALITY?	(Yes if median is in 95% CI for mean)	No	No	No	No
Graphical	Histogram	No	No	No	No
	Box Plot	No	No	No	No
NORMALITY?	(Yes if both are yes, 2 nears=No)	No	No	No	No
Measures of dispersion	Skewness**	3.05	4.451	2.416	2.652
	(within ±0.5)	No	No	No	No
	Kurtosis**	14.97	25.705	7.181	8.143
	(within ±1.0)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
Statistical tests					
Kolmogorov-Smirnov	level of significance	<0.0001	<0.0001	<0.0001	<0.0001
	(not normal if p<0.05)	No	No	No	No
Shapiro-Wilks	level of significance	<0.0001	<0.0001	<0.0001	<0.0001
	(not normal if p<0.05)	No	No	No	No
NORMALITY?	(Yes if both are yes)	No	No	No	No
OVERALL ASSESSMENT OF NORMALITY Yes if there are at least 3 Yeses		**No**	**No**	**No**	**No**

* CI = Confidence Interval

** Z–score of skewness (and kurtosis) were provided by SPSS

Appendix I. Safety ratings for each scene in Experiment 2

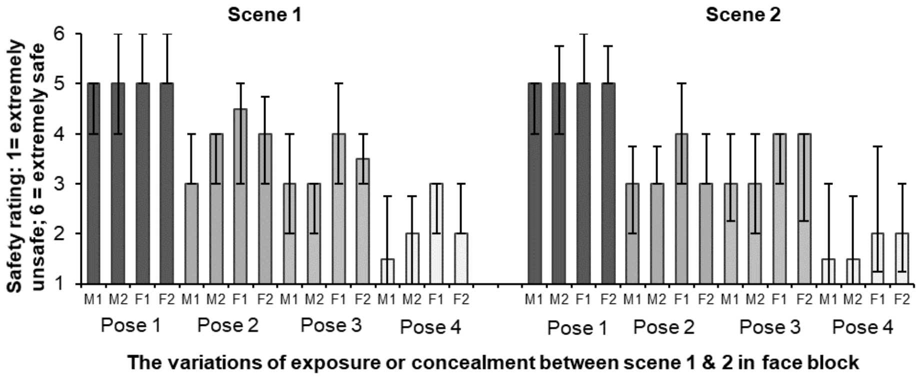


Figure I.1. Median ratings of safety in each background scene, shows each actor in a pose for face block. Error bars show the interquartile range. The poses were illustrated in Table 3.8.

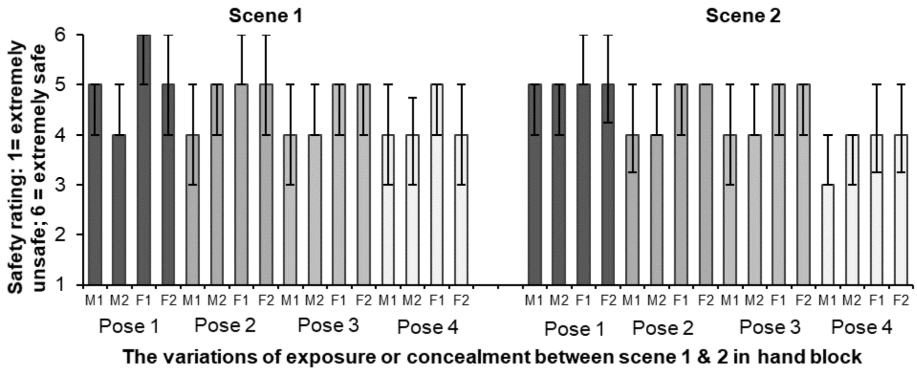


Figure I.2. Median ratings of safety in each background scene, shows each actor in a pose for hands block. Error bars show the interquartile range. The poses were illustrated in Table 3.8.

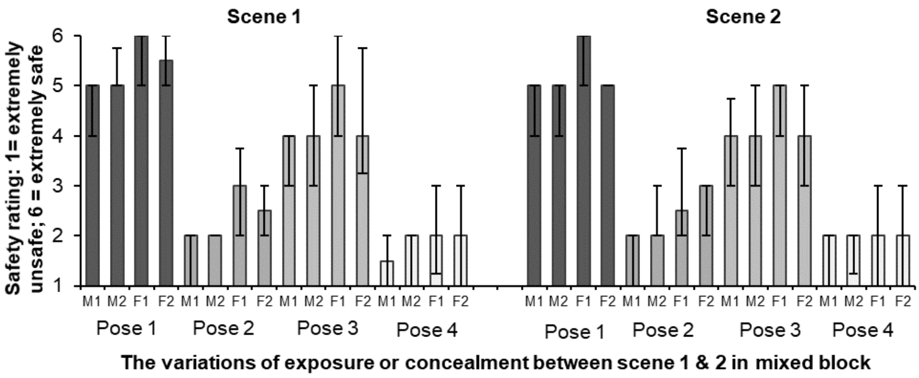


Figure I.3. Median ratings of safety in each background scene, shows each actor in a pose for mixed block. Error bars show the interquartile range. The poses were illustrated in Table 3.8.

Appendix J. Test of significant differences between poses for each actor in Experiment 2

Table J.1. Results of significant differences and the corresponding effect size for each actor in all possible pair comparisons between test poses in face block.

Combination 1							
Pair	**Pose 1**	**Pose 2**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 2	M1	0.82	Large	<0.001	Yes
2	Image 5	Image 6	M2	0.82	Large	<0.001	Yes
3	Image 9	Image 10	F1	0.76	Large	<0.001	Yes
4	Image 13	Image 14	F2	0.80	Large	<0.001	Yes
Combination 2							
Pair	**Pose 1**	**Pose 3**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 3	M1	0.76	Large	<0.001	Yes
2	Image 5	Image 7	M2	0.85	Large	<0.001	Yes
3	Image 9	Image 11	F1	0.75	Large	<0.001	Yes
4	Image 13	Image 15	F2	0.80	Large	<0.001	Yes
Combination 3							
Pair	**Pose 1**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 4	M1	0.85	Large	<0.001	Yes
2	Image 5	Image 8	M2	0.87	Large	<0.001	Yes
3	Image 9	Image 12	F1	0.83	Large	<0.001	Yes
4	Image 13	Image 16	F2	0.86	Large	<0.001	Yes
Combination 4							
Pair	**Pose 2**	**Pose 3**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 2	Image 3	M1	0.06	Non	0.704	No
2	Image 6	Image 7	M2	0.25	Small	0.091	No
3	Image 10	Image 11	F1	0.10	Small	0.493	No
4	Image 14	Image 15	F2	0.09	Non	0.531	No
Combination 5							
Pair	**Pose 2**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 2	Image 4	M1	0.73	Large	<0.001	Yes
2	Image 6	Image 8	M2	0.73	Large	<0.001	Yes
3	Image 10	Image 12	F1	0.72	Large	<0.001	Yes
4	Image 14	Image 16	F2	0.84	Large	<0.001	Yes
Combination 6							
Pair	**Pose 3**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 3	Image 4	M1	0.73	Large	<0.001	Yes
2	Image 7	Image 8	M2	0.59	Large	<0.001	Yes
3	Image 11	Image 12	F1	0.73	Large	<0.001	Yes
4	Image 15	Image 16	F2	0.77	Large	<0.001	Yes

Table J.2. Results of significant differences and the corresponding effect size for each actor in all possible pair comparisons between test poses in hands block.

Combination 1							
Pair	**Pose 1**	**Pose 2**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 2	M1	0.30	Medium	0.046	Yes
2	Image 5	Image 6	M2	0.14	Small	0.346	No
3	Image 9	Image 10	F1	0.30	Medium	0.044	Yes
4	Image 13	Image 14	F2	0.18	Small	0.243	No
Combination 2							
Pair	**Pose 1**	**Pose 3**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 3	M1	0.33	Medium	0.027	Yes
2	Image 5	Image 7	M2	0.20	Small	0.195	No
3	Image 9	Image 11	F1	0.51	large	=0.001	Yes
4	Image 13	Image 15	F2	0.44	Medium	0.003	Yes
Combination 3							
Pair	**Pose 1**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 4	M1	0.56	large	<0.001	Yes
2	Image 5	Image 8	M2	0.57	large	<0.001	Yes
3	Image 9	Image 12	F1	0.56	large	<0.001	Yes
4	Image 13	Image 16	F2	0.62	large	<0.001	Yes
Combination 4							
Pair	**Pose 2**	**Pose 3**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 2	Image 3	M1	0.03	Non	0.827	No
2	Image 6	Image 7	M2	0.02	Non	0.887	No
3	Image 10	Image 11	F1	0.20	Small	0.176	No
4	Image 14	Image 15	F2	0.28	Small	0.066	No
Combination 5							
Pair	**Pose 2**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 2	Image 4	M1	0.46	Medium	0.002	Yes
2	Image 6	Image 8	M2	0.52	large	=0.001	Yes
3	Image 10	Image 12	F1	0.46	Medium	0.002	Yes
4	Image 14	Image 16	F2	0.55	large	<0.001	Yes
Combination 6							
Pair	**Pose 3**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 3	Image 4	M1	0.49	Medium	=0.001	Yes
2	Image 7	Image 8	M2	0.62	large	<0.001	Yes
3	Image 11	Image 12	F1	0.28	Small	0.066	No
4	Image 15	Image 16	F2	0.40	Medium	0.007	Yes

Table J.3. Results of significant differences and the corresponding effect size for each actor in all possible pair comparisons between test poses in mixed block.

Combination 1							
Pair	**Pose 1**	**Pose 2**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 2	M1	0.84	large	<0.001	Yes
2	Image 5	Image 6	M2	0.86	large	<0.001	Yes
3	Image 9	Image 10	F1	0.86	large	<0.001	Yes
4	Image 13	Image 14	F2	0.86	large	<0.001	Yes
Combination 2							
Pair	**Pose 1**	**Pose 3**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 3	M1	0.58	large	<0.001	Yes
2	Image 5	Image 7	M2	0.56	large	<0.001	Yes
3	Image 9	Image 11	F1	0.64	large	<0.001	Yes
4	Image 13	Image 15	F2	0.57	large	<0.001	Yes
Combination 3							
Pair	**Pose 1**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 1	Image 4	M1	0.85	large	<0.001	Yes
2	Image 5	Image 8	M2	0.87	large	<0.001	Yes
3	Image 9	Image 12	F1	0.86	large	<0.001	Yes
4	Image 13	Image 16	F2	0.86	large	<0.001	Yes
Combination 4							
Pair	**Pose 2**	**Pose 3**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 2	Image 3	M1	0.84	large	<0.001	Yes
2	Image 6	Image 7	M2	0.84	large	<0.001	Yes
3	Image 10	Image 11	F1	0.84	large	<0.001	Yes
4	Image 14	Image 15	F2	0.77	large	<0.001	Yes
Combination 5							
Pair	**Pose 2**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 2	Image 4	M1	0.23	Small	0.127	No
2	Image 6	Image 8	M2	0.29	Small	0.052	No
3	Image 10	Image 12	F1	0.46	Medium	0.002	Yes
4	Image 14	Image 16	F2	0.49	Medium	=0.001	Yes
Combination 6							
Pair	**Pose 3**	**Pose 4**	**Actor**	**Effect sizes (r)**	**Effect sizes**	**Wilcoxon test (*p*-value)**	**Sig.**
1	Image 3	Image 4	M1	0.83	large	<0.001	Yes
2	Image 7	Image 8	M2	0.86	large	<0.001	Yes
3	Image 11	Image 12	F1	0.86	large	<0.001	Yes
4	Image 15	Image 16	F2	0.84	large	<0.001	Yes

Appendix K. Calculation of Dunn-Rankin Variance Stable Rank Sums (VSRS) in Experiment 2

Table K.1. Dunn-Rankin VSRS for face block, shows matrix of rank differences between items sums with the significant values.

Face part obscured		**None** (Pose1)	**Top** (Pose2)	**Bottom** (Pose3)	**All** (Pose4)
	Sum	122.4	64.1	69.1	8.5
None (Pose1)	122.4	-	58.3*	53.3*	113.9*
Top (Pose2)	64.1	-	-	5.0	55.6*
Bottom (Pose3)	69.1	-	-	-	60.6*

* Significant at the 0.01 level (critical range = 37.1). The significant difference between items has been identified based on the difference between item sums equal to or larger than the critical range indicates a significant difference (p<0.01).

Table K.2. Dunn-Rankin VSRS for hands block, shows matrix of rank differences between items sums with the significant values.

Hands position		**Side** (Pose1)	**Front** (Pose2)	**Pocket** (Pose3)	**Behind** (Pose4)
	Sum	94.6	77.9	67.0	24.4
Side (Pose1)	94.6	-	16.7	27.6	70.2*
Front (Pose2)	77.9	-	-	10.9	53.5*
Pocket (Pose3)	67.0	-	-	-	42.6*

* Significant at the 0.01 level (critical range = 37.1). The significant difference between items has been identified based on the difference between item sums equal to or larger than the critical range indicates a significant difference (p<0.01).

Table K.3. Dunn-Rankin VSRS for mixed block, shows matrix of rank differences between items sums with the significant values.

Mixed condition		**Face and hands exposed** (Pose1)	**Face concealed hands exposed** (Pose2)	**Face exposed hands concealed** (Pose3)	**Face and hands concealed** (Pose4)
	Sum	120.9	42.3	88.8	12.0
Face and hands exposed (Pose1)	120.9	-	78.6*	32.1*	108.9*
Face concealed hands exposed (Pose2)	42.3	-	-	46.5*	30.3
Face exposed hands concealed (Pose3)	88.8	-	-	-	76.8*

* Significant at the 0.01 level (critical range = 37.1). The significant difference between items has been identified based on the difference between item sums equal to or larger than the critical range indicates a significant difference (p<0.01).

Appendix L. The images used to generate the test images for Experiment 3

Figure L.1. The 20 images of the target person, shows the 10 males and the 10 females.

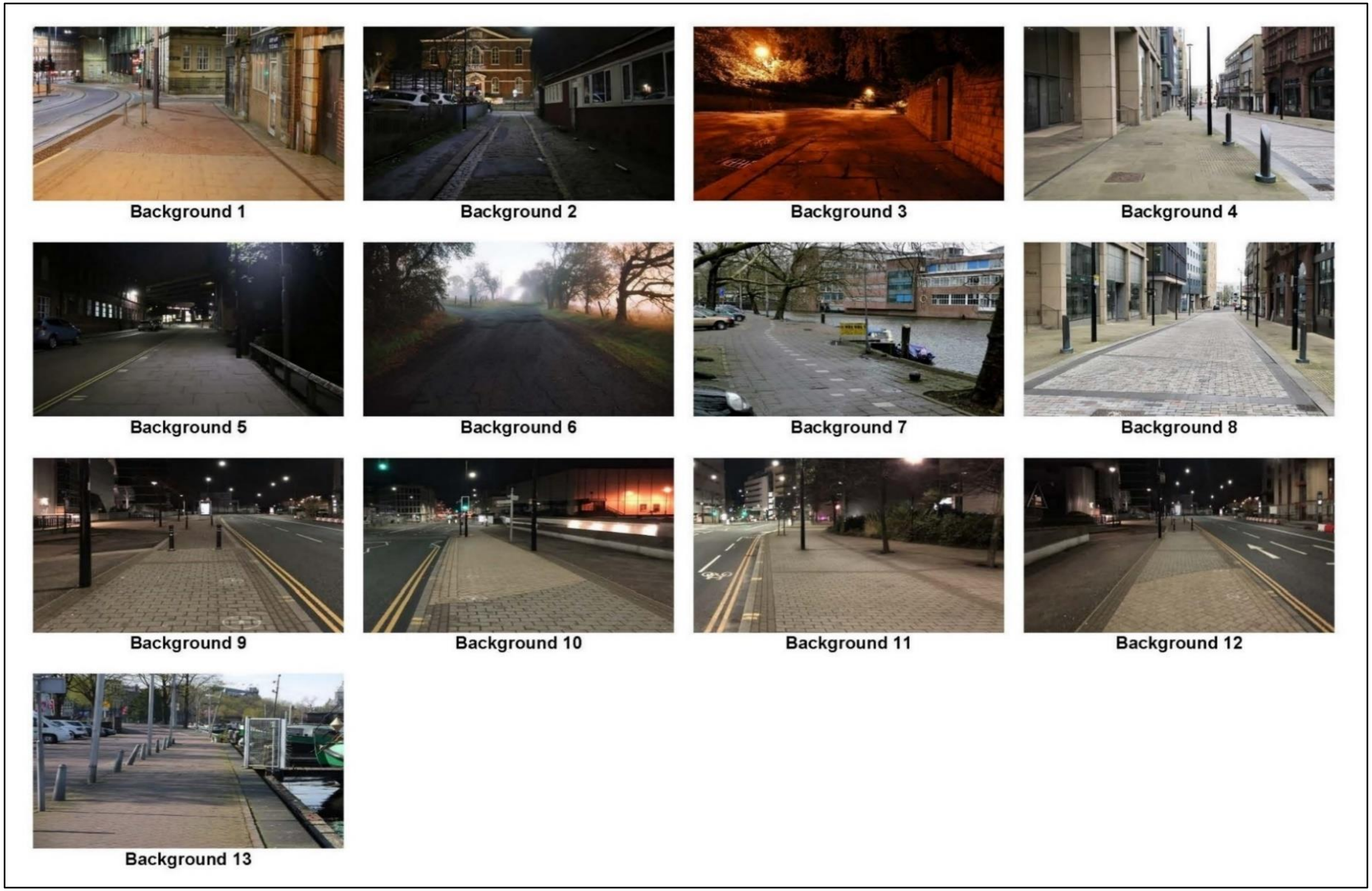


Figure L.2. The 13 images of the background scenes.

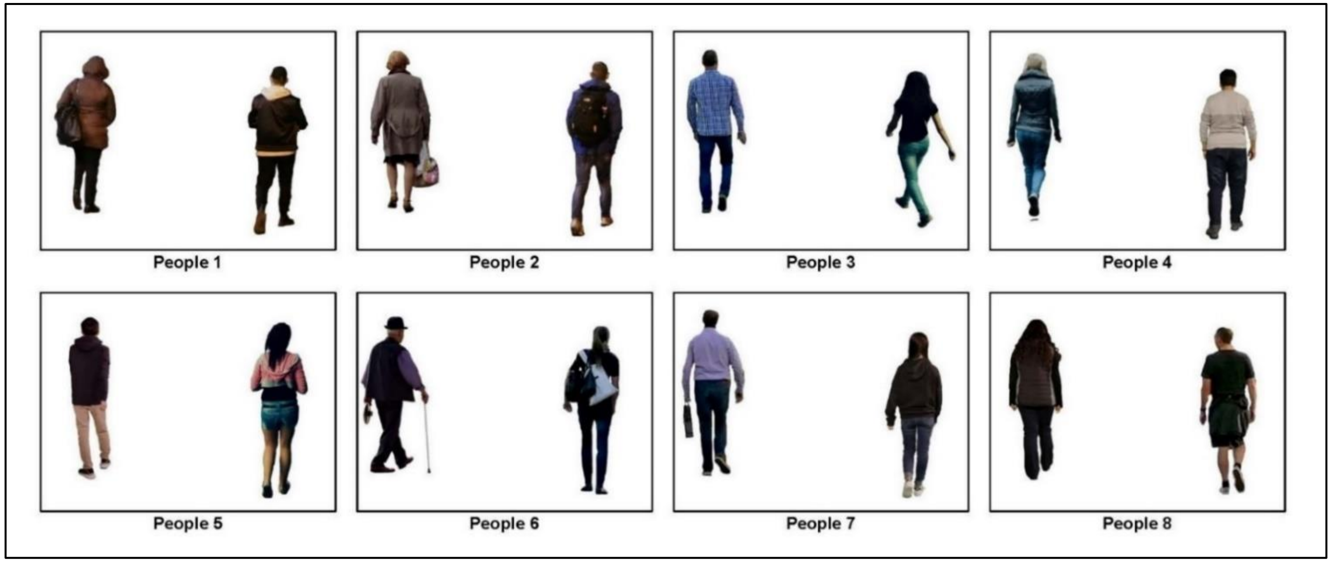


Figure L.3. The eight images of the pair of people.

Figure L.4. The six images of the target person with a dog.

Figure L.5. The five background images with vehicle(s).

Appendix M. The images and the form used for the validation exercise in Experiment 3

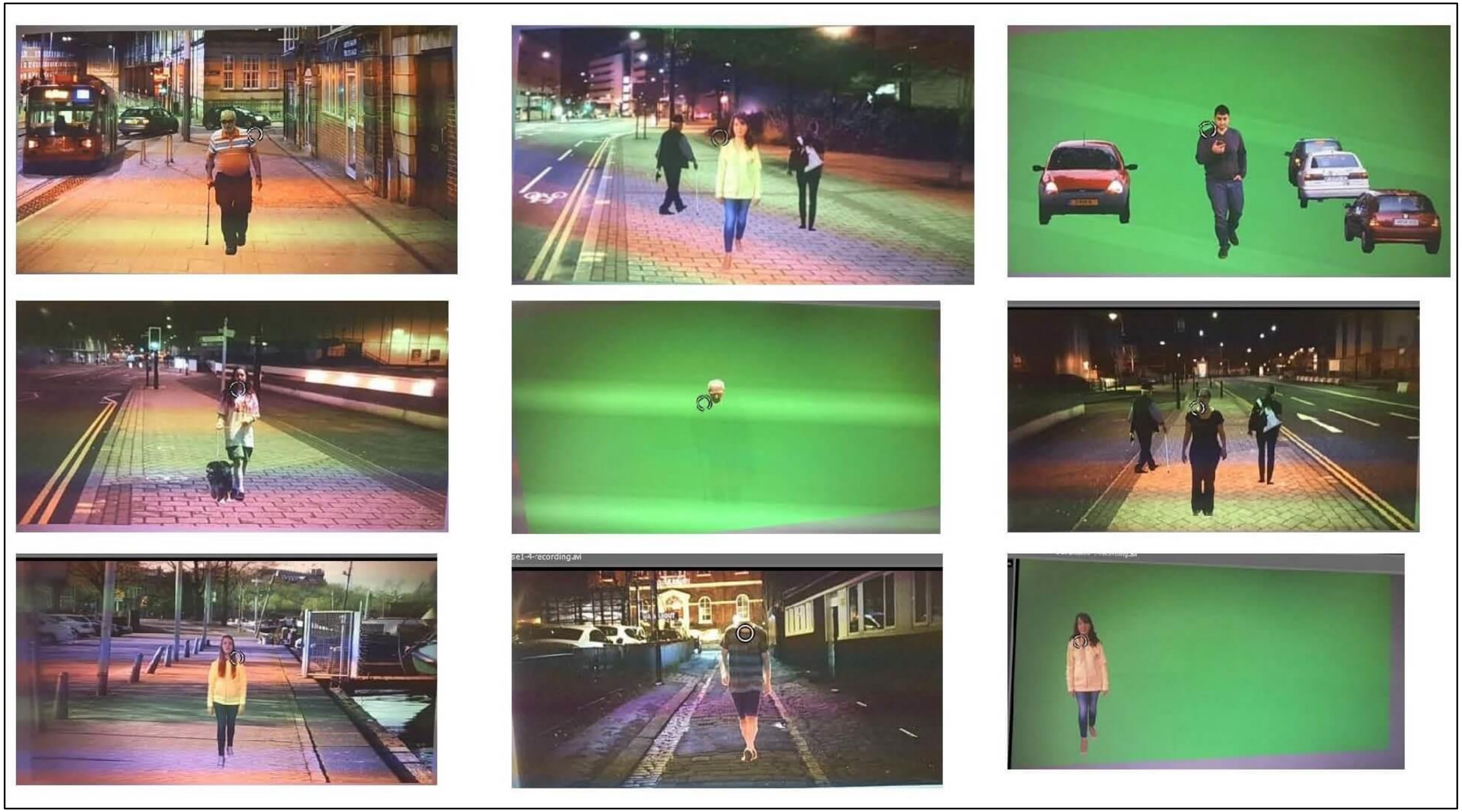


Figure M.1. The nine images where the experimenter was unsure about the fixation.

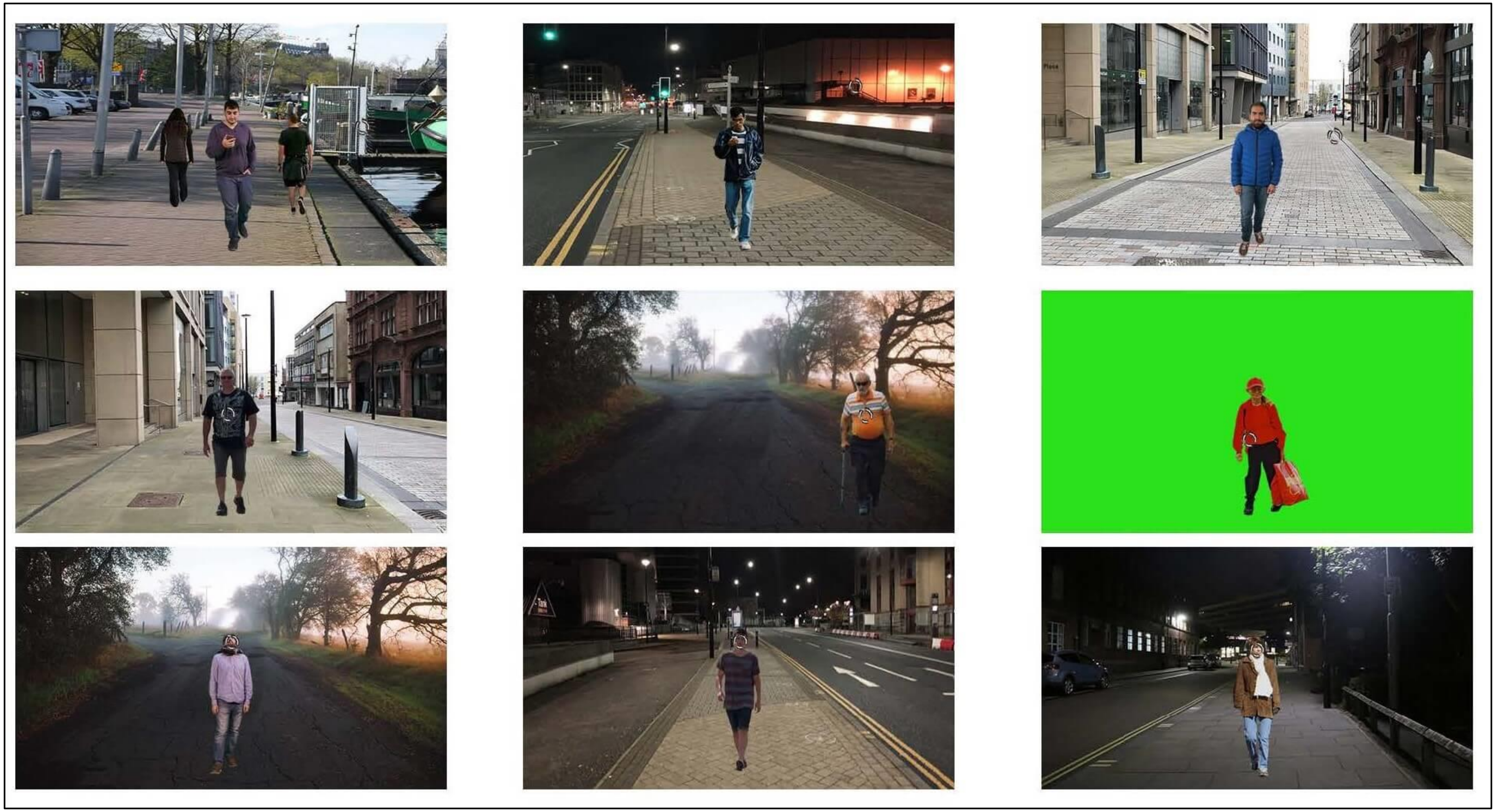


Figure M.2. The nine images where the experimenter was confident about the fixation domain.

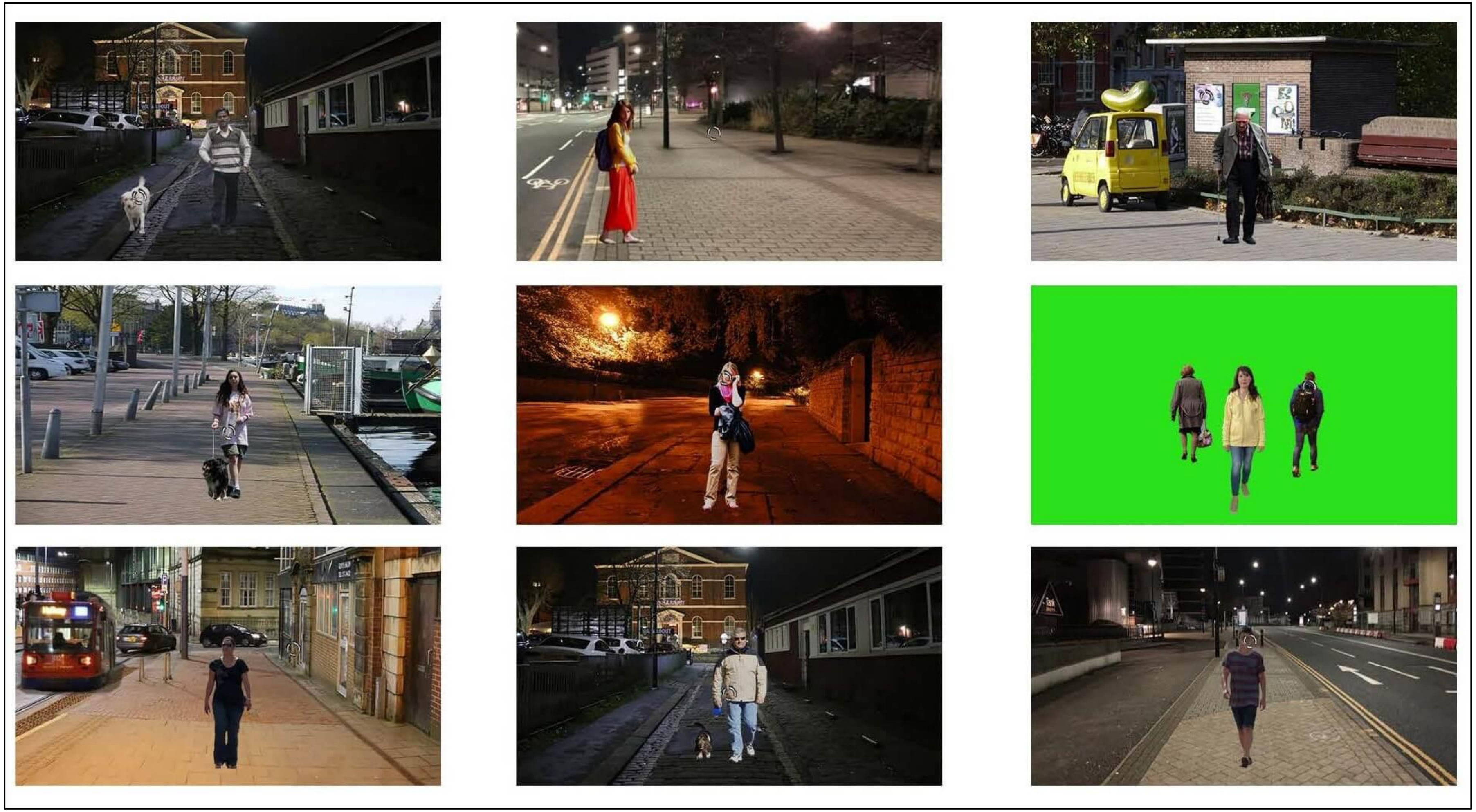
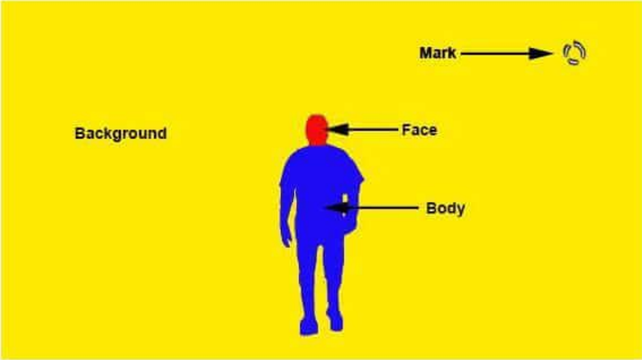


Figure M.3. The nine images that were selected randomly.

Validation of fixation allocation

Validation number: **Date:**

In the following images, please circle the relevant keyword below the image (face, body or background, other) to show which element of the image is the focus of visual fixation. Where the choice is not obvious, please pick the *one* element you think is most likely.



An example illustrates the elements' area.

Face **Body** **Background**

Other: dog, people, vehicle

Face **Body** **Background**

Other: dog, people, vehicle

Face **Body** **Background**

Other: dog, people, vehicle

Face **Body** **Background**

Other: dog, people, vehicle

Face **Body** **Background**

Other: dog, people, vehicle

Face **Body** **Background**

Other: dog, people, vehicle

Face **Body** **Background**

Other: dog, people, vehicle

Face **Body** **Background**

Other: dog, people, vehicle

Face **Body** **Background**

Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

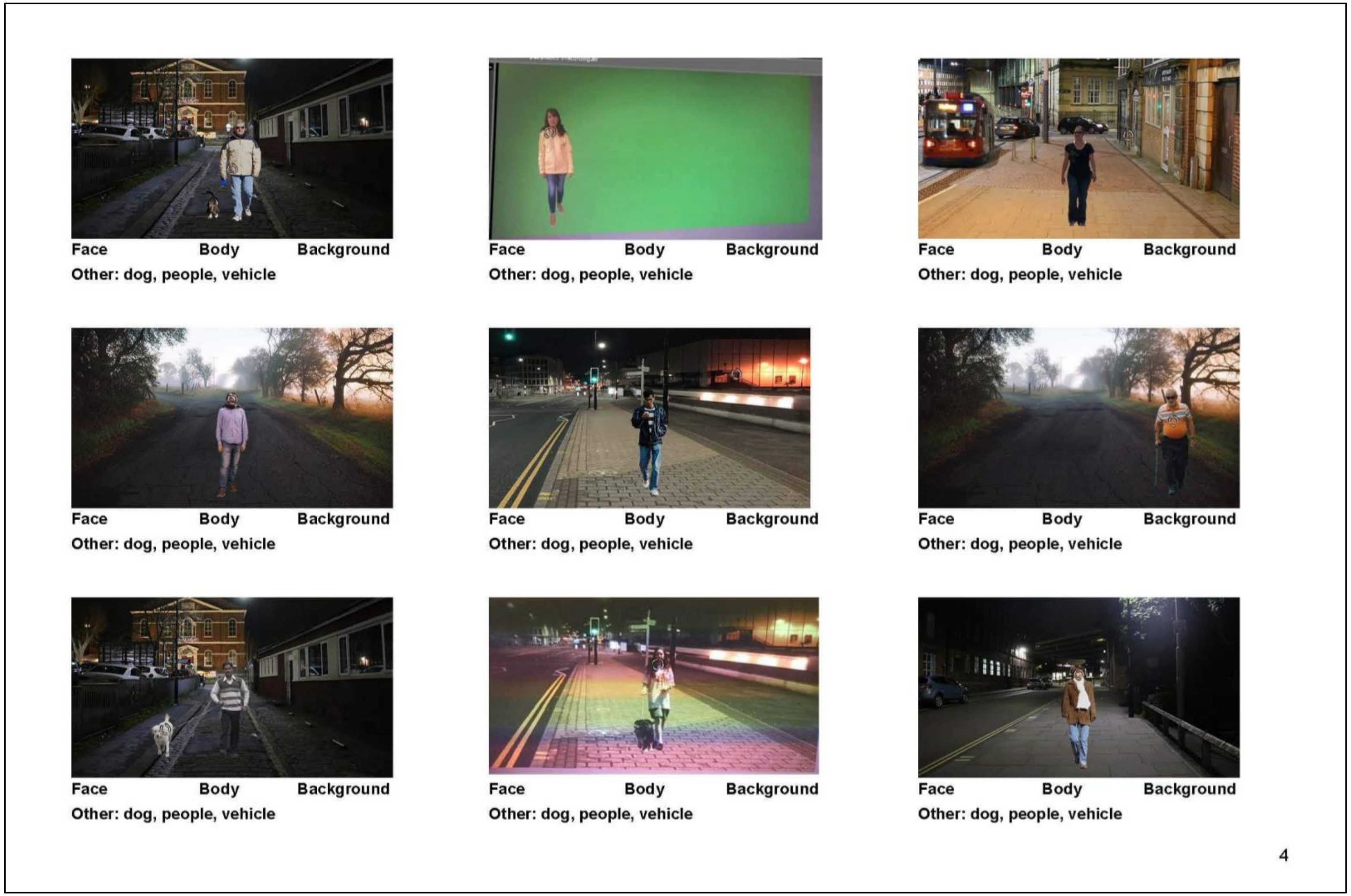
Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle



Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

Face Body Background
Other: dog, people, vehicle

4

Figure M.4. The form used for the validation exercise in Experiment 3.

Appendix N. Results of fixations for male and female targets and across both in Experiment 3

Table N.1. Results of normal-size scores for the fixation durations for male and female targets and across both for each test image in Block 1. Note: the mean unit is millisecond.

Category of test images	Gender	n	Face			Body			Background			Distraction		
			Mean	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**
Primary	F	14	1065	970	42%	553	621	22%	925	720	36%	---	---	---
	M	18	966	757	41%	685	930	29%	708	676	30%	---	---	---
	F&M	32	1009	843	41%	627	800	26%	803	692	33%	---	---	---
Left position	F	22	425	557	18%	582	654	25%	1340	597	57%	---	---	---
	M	26	473	612	19%	651	512	26%	1337	711	54%	---	---	---
	F&M	48	451	582	19%	619	576	26%	1338	655	56%	---	---	---
Right position	F	22	428	421	17%	293	430	12%	1760	693	71%	---	---	---
	M	26	593	591	25%	477	527	20%	1325	635	55%	---	---	---
	F&M	48	518	521	21%	392	489	16%	1525	691	63%	---	---	---
Dog	F	19	523	551	21%	721	501	28%	716	544	28%	572	569	23%
	M	13	800	925	31%	486	463	19%	863	760	33%	471	406	18%
	F&M	32	635	726	25%	626	492	24%	776	633	30%	531	504	21%
People	F	16	492	500	21%	339	549	14%	826	517	35%	707	466	30%
	M	16	886	835	36%	424	481	17%	676	634	28%	460	444	19%
	F&M	32	689	706	29%	382	509	16%	751	574	31%	583	465	24%
Vehicles	F	16	712	607	31%	537	596	23%	332	408	14%	719	479	31%
	M	16	1118	732	44%	328	416	13%	479	395	19%	608	753	24%
	F&M	32	915	693	38%	432	516	18%	406	402	17%	663	623	27%
Obscured face	F	23	601	775	26%	990	746	42%	756	520	32%	---	---	---
	M	24	521	650	21%	1030	708	42%	873	590	36%	---	---	---
	F&M	47	560	707	23%	1010	719	42%	816	554	34%	---	---	---
Obscured body	F	28	1086	835	45%	632	799	26%	722	565	30%	---	---	---
	M	21	1117	877	42%	592	760	22%	949	696	36%	---	---	---
	F&M	49	1099	844	43%	615	775	24%	819	628	32%	---	---	---
All images	F	160	669	709	23%	595	658	21%	943	701	33%	660*	505*	23%*
	M	160	772	759	27%	607	648	21%	953	696	33%	516**	557**	18%**
	F&M	320	721	735	25%	601	652	21%	948	698	33%	593***	532***	21%***

* n = 51, ** n = 54, *** n = 96

Table N.2. Results of area-weighted scores for the fixation durations for male and female targets and across both for each test image in Block 1. Note: the mean unit is millisecond.

Category of test images	**Gender**	**n**	**Face**			**Body**			**Background**			**Distraction**		
			Mean	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**
Primary	F	14	253778	231024	95%	11846	13299	4%	974	758	0.4%	---	---	---
	M	18	230179	180337	94%	14688	19942	6%	746	712	0.3%	---	---	---
	F&M	32	240504	200892	94%	13445	17155	5%	846	730	0.3%	---	---	---
Left position	F	22	101283	132762	88%	12483	14014	11%	1411	629	1%	---	---	---
	M	26	112693	145910	88%	13950	10976	11%	1409	749	1%	---	---	---
	F&M	48	107463	138682	88%	13277	12344	11%	1410	690	1%	---	---	---
Right position	F	22	102004	100265	93%	6273	9213	6%	1854	730	2%	---	---	---
	M	26	141375	140776	92%	10218	11303	7%	1396	669	1%	---	---	---
	F&M	48	123330	124202	92%	8410	10480	6%	1606	728	1%	---	---	---
Dog	F	19	124589	131393	62%	15447	10735	8%	763	579	0.4%	59219	58873	30%
	M	13	190581	220333	76%	10424	9919	4%	918	808	0.4%	48781	42036	19%
	F&M	32	151398	172919	69%	13406	10549	6%	826	674	0.4%	54979	52189	25%
People	F	16	117136	119253	82%	7266	11762	5%	911	570	1%	16709	11006	12%
	M	16	211109	198914	91%	9094	10310	4%	745	699	0.3%	10859	10497	5%
	F&M	32	164122	168242	88%	8180	10919	4%	828	633	0.4%	13784	10989	7%
Vehicles	F	16	169536	144671	92%	11502	12765	6%	447	550	0.2%	3463	2306	2%
	M	16	266390	174345	96%	7020	8907	3%	646	532	0.2%	2930	3627	1%
	F&M	32	217963	165094	94%	9261	11065	4%	547	542	0.2%	3196	3002	1%
Obscured face	F	23	143237	184725	87%	21210	15987	13%	797	548	0.5%	---	---	---
	M	24	124055	154794	84%	22066	15171	15%	920	621	1%	---	---	---
	F&M	47	133442	168506	86%	21647	15411	14%	859	584	1%	---	---	---
Obscured body	F	28	258787	198900	95%	13540	17122	5%	761	595	0.3%	---	---	---
	M	21	266074	209005	95%	12686	16291	5%	1000	734	0.4%	---	---	---
	F&M	49	261910	201166	95%	13174	16603	5%	863	662	0.3%	---	---	---
All images	F	160	159498	168964	79%	12745	14094	6%	1008	740	1%	28390*	43467*	14%*
	M	160	183946	180909	85%	13013	13892	6%	1023	734	0.5%	18995**	30067**	9%**
	F&M	320	171722	175193	82%	12879	13972	6%	1016	736	0.5%	23986***	37886***	11%***

* n = 51, ** n = 54, *** n = 96

Table N.3. Results of normal-size scores for the fixation durations for male and female targets and across both for each test image in Block 2. Note: the mean unit is millisecond.

Category of test images	Gender	n	Face			Body			Background			Distraction		
			Mean	Stdev.	%	Mean	Stdev.	%	Mean	Stdev.	%	Mean	Stdev.	%
Primary	F	16	855	686	36%	625	645	27%	863	593	37%	---	---	---
	M	16	1263	865	52%	411	649	17%	736	702	31%	---	---	---
	F&M	32	1059	795	45%	518	646	22%	800	642	34%	---	---	---
Left position	F	20	560	527	25%	478	507	22%	1180	568	53%	---	---	---
	M	24	599	650	29%	578	579	28%	911	673	44%	---	---	---
	F&M	43	586	597	27%	539	548	25%	1015	632	47%	---	---	---
Right position	F	28	764	591	33%	283	376	12%	1281	643	55%	---	---	---
	M	24	762	633	32%	361	375	15%	1256	723	53%	---	---	---
	F&M	54	1131	2872	33%	454	1099	13%	1887	4761	54%	---	---	---
Dog	F	20	951	458	40%	383	364	16%	660	504	28%	394	329	16%
	M	12	682	726	29%	480	497	20%	904	599	38%	287	284	12%
	F&M	32	850	577	36%	419	414	18%	751	546	32%	354	312	15%
People	F	14	774	481	35%	209	207	9%	571	568	26%	655	458	30%
	M	18	724	512	29%	513	559	21%	431	344	17%	817	341	33%
	F&M	32	746	491	32%	380	461	16%	492	453	21%	746	398	32%
Vehicles	F	17	1104	736	49%	260	309	12%	353	444	16%	532	370	24%
	M	15	995	647	39%	528	450	21%	313	357	12%	687	256	27%
	F&M	34	1543	3121	46%	493	798	15%	491	1050	15%	834	1487	25%
Obscured face	F	27	638	635	27%	1014	723	42%	735	499	31%	---	---	---
	M	21	564	525	24%	995	632	42%	788	665	34%	---	---	---
	F&M	48	606	585	26%	1006	678	42%	758	571	32%	---	---	---
Obscured body	F	23	1215.5	852.0	52%	256	461	11%	877	667	37%	---	---	---
	M	25	1195	689	51%	363	439	16%	777	528	33%	---	---	---
	F&M	48	1205	763	51%	312	448	13%	825	594	35%	---	---	---
All images	F	168	1070	2199	32%	523	852	16%	1075	2801	32%	662*	1235*	20%*
	M	155	843	689	30%	529	552	19%	794	648	28%	632**	365**	23%**
	F&M	323	961	1657	31%	526	723	17%	940	2071	31%	649***	937***	21%***

* n = 53, ** n = 45, *** n = 98

Table N.4. Results of area-weighted scores for the fixation durations for male and female targets and across both for each test image in Block 2. Note: the mean unit is millisecond.

Category of test images	n	Gender	Face			Body			Background			Distraction		
			Mean	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**	**Mean**	**Stdev.**	**%**
Primary	16	F	203715	163434	93%	13395	13822	6%	909	625	0.4%	---	---	---
	16	M	300844	206049	97%	8804	13905	3%	776	739	0.2%	---	---	---
	32	F&M	252280	189479	95%	11100	13836	4%	843	677	0.3%	---	---	---
Left position	20	F	133429	125513	92%	10254	10864	7%	1243	599	1%	---	---	---
	24	M	142732	154809	91%	12379	12403	8%	960	709	1%	---	---	---
	43	F&M	139512	142223	92%	11562	11739	8%	1070	665	1%	---	---	---
Right position	28	F	181973	140784	96%	6063	8050	3%	1350	677	1%	---	---	---
	24	M	181549	150792	95%	7727	8031	4%	1324	762	1%	---	---	---
	54	F&M	269401	684309	96%	9722	23558	3%	1988	5016	1%	---	---	---
Dog	20	F	226586	109103	82%	8202	7801	3%	702	537	0.3%	40754	34024	15%
	12	M	162571	172910	80%	10282	10654	5%	962	637	0.5%	29720	29440	15%
	32	F&M	202581	137463	81%	8982	8867	4%	800	581	0.3%	36616	32349	15%
People	14	F	184377	114565	90%	4483	4430	2%	629	627	0.3%	15472	10818	8%
	18	M	172478	122090	85%	10996	11985	5%	475	380	0.2%	19307	8057	9%
	32	F&M	177684	117108	87%	8147	9888	4%	543	500	0.3%	17629	9403	9%
Vehicles	17	F	263070	175289	97%	5580	6617	2%	476	599	0.2%	2561	1784	1%
	15	M	237084	154109	94%	11318	9644	4%	422	482	0.2%	3308	1233	1%
	34	F&M	367665	743621	96%	10573	17101	3%	663	1416	0.2%	4020	7164	1%
Obscured face	27	F	152028	151310	87%	21729	15496	12%	775	526	0.4%	---	---	---
	21	M	134434	125167	86%	21331	13543	14%	831	701	1%	---	---	---
	48	F&M	144331	139318	87%	21555	14523	13%	799	602	0.5%	---	---	---
Obscured body	23	F	289622	203015	98%	5489	9887	2%	924	702	0.3%	---	---	---
	25	M	284826	164148	97%	7774	9413	3%	818	557	0.3%	---	---	---
	48	F&M	287124	181815	97%	6679	9609	2%	869	626	0.3%	---	---	---
All images	168	F	255041	523865	88%	11211	18254	4%	1157	2972	0.4%	21108*	27220*	7%*
	155	M	200979	164287	87%	11334	11838	5%	849	683	0.4%	16751**	18772**	7%**
	323	F&M	229098	394932	88%	11270	15487	4%	1009	2197	0.4%	19108***	23702***	7%***

* n = 53, ** n = 45, *** n = 98

Table N.5. Results of male and female targets and across both, show the significant differences using t-test (p = 0.05) and Bonferroni corrections (p = 0.02, p = 0.009) (shaded) between test domains in each category of test images; this applied to the fixation durations for the normal-size and area-weighted scores in Block 1.

Category of test images	**Gender**	**n**	**Normal-size**						**Area-weighted**					
			Face v Body	**Face v Bg.***	**Body v Bg.***	**Dis.** v Face**	**Dis.** v Body**	**Dis.** v Bg.***	**Face v Body**	**Face v Bg.***	**Body v Bg.***	**Dis.** v Face**	**Dis.** v Body**	**Dis.** v Bg.***
Primary	F	14	0.185	0.732	0.244	---	---	---	0.002	0.001	0.011	---	---	---
	M	18	0.428	0.000	0.010	---	---	---	0.000	0.000	0.010	---	---	---
	F&M	32	0.000	0.000	0.000	---	---	---	0.000	0.000	0.000	---	---	---
Left position	F	22	0.449	0.000	0.005	---	---	---	0.006	0.002	0.002	---	---	---
	M	26	0.311	0.001	0.003	---	---	---	0.002	0.001	0.000	---	---	---
	F&M	48	0.203	0.000	0.000	---	---	---	0.000	0.000	0.000	---	---	---
Right position	F	22	0.384	0.000	0.000	---	---	---	0.000	0.000	0.040	---	---	---
	M	26	0.526	0.001	0.000	---	---	---	0.000	0.000	0.001	---	---	---
	F&M	48	0.299	0.000	0.000	---	---	---	0.000	0.000	0.000	---	---	---
Dog	F	19	0.359	0.328	0.983	0.815	0.397	0.527	0.003	0.001	0.000	0.090	0.005	0.000
	M	13	0.367	0.880	0.219	0.328	0.913	0.140	0.014	0.009	0.006	0.056	0.004	0.001
	F&M	32	0.958	0.483	0.172	0.565	0.417	0.149	0.000	0.000	0.000	0.009	0.000	0.000
People	F	16	0.510	0.150	0.023	0.083	0.091	0.584	0.003	0.001	0.048	0.008	0.002	0.000
	M	16	0.121	0.514	0.285	0.105	0.835	0.335	0.001	0.001	0.006	0.001	0.645	0.002
	F&M	32	0.097	0.752	0.018	0.476	0.139	0.271	0.000	0.000	0.001	0.000	0.070	0.000
Vehicles	F	16	0.518	0.061	0.336	0.970	0.419	0.024	0.001	0.000	0.004	0.000	0.033	0.000
	M	16	0.003	0.015	0.362	0.151	0.252	0.588	0.000	0.000	0.013	0.000	0.136	0.029
	F&M	32	0.011	0.002	0.843	0.210	0.154	0.073	0.000	0.000	0.000	0.000	0.008	0.000
Obscured face	F	23	0.191	0.453	0.312	---	---	---	0.007	0.453	0.000	---	---	---
	M	24	0.035	0.115	0.492	---	---	---	0.005	0.115	0.000	---	---	---
	F&M	47	0.017	0.089	0.224	---	---	---	0.000	0.000	0.000	---	---	---
Obscured body	F	28	0.106	0.084	0.690	---	---	---	0.000	0.000	0.001	---	---	---
	M	21	0.102	0.586	0.199	---	---	---	0.000	0.000	0.004	---	---	---
	F&M	49	0.020	0.112	0.238	---	---	---	0.000	0.000	0.000	---	---	---

* Bg.= Background.
** Dis.= Distraction(s), this could be either dogs, people, or vehicles, as shown in the subcategory.
Shaded= significant using Bonferroni corrections.

Table N.6. Results of male and female targets and across both, show the significant differences using t-test (p = 0.05) and Bonferroni corrections (p = 0.02, p = 0.009) (shaded) between test domains in each category of test images; this applied to the fixation durations for the normal-size and area-weighted scores in Block 2.

Category of test images	**Gender**	**n**	**Normal-size**						**Area-weighted**					
			Face v Body	**Face v Bg.***	**Body v Bg.***	**Dis.** v Face**	**Dis.** v Body**	**Dis.** v Bg.***	**Face v Body**	**Face v Bg.***	**Body v Bg.***	**Dis.** v Face**	**Dis.** v Body**	**Dis.** v Bg.***
Primary	F	16	0.423	0.976	0.394	---	---	---	0.000	0.000	0.003	---	---	---
	M	16	0.016	0.129	0.263	---	---	---	0.000	0.000	0.039	---	---	---
	F&M	32	0.017	0.233	0.152	---	---	---	0.000	0.000	0.000	---	---	---
Left position	F	20	0.670	0.006	0.001	---	---	---	0.000	0.000	0.002	---	---	---
	M	24	0.916	0.1441	0.092	---	---	---	0.001	0.000	0.000	---	---	---
	F&M	43	0.746	0.005	0.001	---	---	---	0.000	0.000	0.000	---	---	---
Right position	F	28	0.003	0.010	0.000	---	---	---	0.000	0.000	0.006	---	---	---
	M	24	0.011	0.064	0.000	---	---	---	0.000	0.000	0.001	---	---	---
	F&M	54	0.012	0.014	0.007	---	---	---	0.006	0.006	0.004	---	---	---
Dog	F	20	0.002	0.146	0.100	0.000	0.903	0.121	0.000	0.000	0.000	0.000	0.000	0.000
	M	12	0.503	0.504	0.143	0.120	0.069	0.018	0.012	0.008	0.013	0.026	0.012	0.006
	F&M	32	0.006	0.572	0.062	0.000	0.334	0.006	0.000	0.000	0.000	0.000	0.000	0.000
People	F	14	0.001	0.371	0.063	0.428	0.006	0.710	0.000	0.000	0.009	0.000	0.001	0.000
	M	18	0.334	0.106	0.659	0.518	0.121	0.005	0.000	0.000	0.002	0.000	0.000	0.000
	F&M	32	0.010	0.068	0.407	0.996	0.004	0.040	0.000	0.000	0.000	0.000	0.001	0.000
Vehicles	F	17	0.001	0.009	0.568	0.017	0.054	0.167	0.000	0.000	0.008	0.000	0.106	0.000
	M	15	0.089	0.012	0.155	0.130	0.326	0.004	0.000	0.000	0.001	0.000	0.009	0.000
	F&M	34	0.020	0.011	0.986	0.027	0.040	0.005	0.007	0.007	0.001	0.007	0.003	0.002
Obscured face	F	27	0.119	0.562	0.156	---	---	---	0.000	0.562	0.000	---	---	---
	M	21	0.044	0.330	0.415	---	---	---	0.001	0.330	0.000	---	---	---
	F&M	48	0.014	0.260	0.109	---	---	---	0.000	0.000	0.000	---	---	---
Obscured body	F	23	0.000	0.236	0.002	---	---	---	0.000	0.000	0.039	---	---	---
	M	25	0.000	0.041	0.008	---	---	---	0.000	0.000	0.001	---	---	---
	F&M	48	0.000	0.026	0.000	---	---	---	0.000	0.000	0.000	---	---	---

* Bg.= Background.
** Dis.= Distraction(s), this could be either dogs, people, or vehicles, as shown in the subcategory.
Shaded= significant using Bonferroni corrections.

Table N.7. Percentage of first fixations and the number of images for male and female targets and across both for each test image; this applied to the normal-size and area-weighted scores in Block 1.

Category of test images	Gender	Normal-size								Area-weighted							
		Face		Body		Background		Distraction		Face		Body		Background		Distraction	
		%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n
Primary	F	50%	4	17%	7	33%	3	---	---	97%	4	3%	7	0.3%	3	---	---
	M	59%	7	12%	6	28%	5	---	---	98%	7	2%	6	0.2%	5	---	---
	F&M	55%	11	15%	13	30%	8	---	---	97%	11	2%	13	0.2%	8	---	---
Left position	F	34%	2	18%	9	48%	11	---	---	95%	2	5%	9	1%	11	---	---
	M	0%	0	46%	11	54%	15	---	---	0%	0	95%	11	5%	15	---	---
	F&M	17%	2	32%	20	51%	26	---	---	47%	2	50%	20	3%	26	---	---
Right position	F	0%	2	30%	2	70%	17	---	---	0%	2	90%	2	10%	17	---	---
	M	44%	9	14%	1	42%	16	---	---	97%	9	3%	1	0.4%	16	---	---
	F&M	22%	11	22%	3	56%	33	---	---	48%	11	46%	3	5%	33	---	---
Dog	F	42%	5	22%	10	36%	4	0%	0	95%	5	4%	10	0.4%	4	0%	0
	M	15%	2	13%	5	61%	5	10%	1	73%	2	6%	5	1%	5	20%	1
	F&M	29%	7	18%	15	49%	9	5%	1	84%	7	5%	15	1%	9	10%	1
Pair of people	F	38%	4	12%	4	24%	7	26%	1	86%	4	2%	4	0.3%	7	12%	1
	M	48%	3	14%	6	25%	5	12%	2	93%	3	2%	6	0.2%	5	5%	2
	F&M	43%	7	13%	10	25%	12	19%	3	89%	7	2%	10	0.2%	12	8%	3
Vehicles	F	22%	5	17%	8	12%	2	49%	1	90%	5	6%	8	0.3%	2	4%	1
	M	51%	4	13%	4	11%	3	25%	5	97%	4	2%	4	0.1%	3	1%	5
	F&M	36%	9	15%	12	12%	5	37%	6	93%	9	4%	12	0.2%	5	3%	6
Obscured face	F	63%	7	24%	9	14%	7	---	---	97%	7	3%	9	0.1%	7	---	---
	M	26%	3	29%	12	45%	9	---	---	90%	3	9%	12	1%	9	---	---
	F&M	44%	10	26%	21	29%	16	---	---	93%	10	6%	21	0%	16	---	---
Obscured body	F	41%	8	42%	6	16%	13	---	---	91%	8	8%	6	0.2%	13	---	---
	M	51%	6	33%	6	16%	9	---	---	94%	6	5%	6	0.1%	9	---	---
	F&M	46%	14	38%	12	16%	22	---	---	93%	14	7%	12	0.1%	22	---	---
All images	F	36%	37	23%	55	32%	64	0	2	81%	37	15%	55	2%	64	2%	2
	M	37%	34	22%	51	35%	67	0	8	80%	34	16%	51	1%	67	3%	8
	F&M	37%	71	22%	106	33%	131	0	10	81%	71	15%	106	1%	131	3%	10

Table N.8. Percentage of first fixations and the number of images for male and female targets and across both for each test image; this applied to the normal-size and area-weighted scores in Block 2.

Category of test images	**Gender**	**Normal-size**								**Area-weighted**							
		Face		**Body**		**Background**		**Distraction**		**Face**		**Body**		**Background**		**Distraction**	
		%	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**
Primary	F	27%	8	40%	5	34%	3	---	---	88%	8	12%	5	0.5%	3	---	---
	M	36%	11	32%	4	31%	1	---	---	92%	11	7%	4	0.4%	1	---	---
	F&M	32%	19	36%	9	33%	4	---	---	90%	19	10%	9	0.4%	4	---	---
Left position	F	52%	5	32%	7	16%	7	---	---	95%	5	5%	7	0.1%	7	---	---
	M	54%	1	22%	11	25%	10	---	---	96%	1	3%	11	0.2%	10	---	---
	F&M	53%	6	27%	18	20%	17	---	---	95%	6	4%	18	0.2%	17	---	---
Right position	F	72%	4	15%	6	13%	18	---	---	98%	4	2%	6	0.1%	18	---	---
	M	40%	7	28%	3	32%	14	---	---	94%	7	6%	3	0.3%	14	---	---
	F&M	56%	11	22%	9	22%	32	---	---	96%	11	4%	9	0.2%	32	---	---
Dog	F	50%	8	19%	7	31%	5	0%	0	96%	8	3%	7	0.3%	5	0%	0
	M	52%	5	30%	5	18%	2	0%	0	95%	5	5%	5	0.2%	2	0%	0
	F&M	51%	13	25%	12	25%	7	0%	0	96%	13	4%	12	0.2%	7	0%	0
Pair of people	F	34%	5	25%	2	21%	4	19%	2	85%	5	6%	2	0.2%	4	9%	2
	M	45%	6	31%	7	0%	0	24%	5	86%	6	5%	7	0.0%	0	9%	5
	F&M	40%	11	28%	9	11%	4	22%	7	85%	11	5%	9	0.1%	4	9%	7
Vehicles	F	27%	10	11%	2	49%	2	13%	3	94%	10	4%	2	1%	2	1%	3
	M	36%	6	43%	5	21%	4	0%	0	90%	6	10%	5	0.3%	4	0%	0
	F&M	31%	16	27%	7	35%	6	6%	3	92%	16	7%	7	1%	6	0%	3
Obscured face	F	37%	2	31%	21	32%	3	---	---	93%	2	7%	21	0.4%	3	---	---
	M	30%	6	53%	11	17%	4	---	---	86%	6	14%	11	0.2%	4	---	---
	F&M	33%	8	42%	32	25%	7	---	---	89%	8	10%	32	0.3%	7	---	---
Obscured body	F	59%	13	14%	3	27%	6	---	---	98%	13	2%	3	0.2%	6	---	---
	M	43%	12	43%	3	14%	8	---	---	92%	12	8%	3	0.1%	8	---	---
	F&M	51%	25	29%	6	20%	14	---	---	95%	25	5%	6	0.2%	14	---	---
All images	F	45%	55	23%	53	28%	48	4%	5	93%	55	5%	53	0.3%	48	1%	5
	M	42%	54	35%	49	20%	43	3%	5	91%	54	7%	49	0.2%	43	1%	5
	F&M	43%	109	29%	102	24%	91	4%	10	92%	109	6%	102	0.3%	91	1%	10

References

Abele, A., & Yzerbyt, V. (2021). Body posture and interpersonal perception in a dyadic interaction: A Big Two analysis. *European Journal of Social Psychology*, *51*(1), 23–39.

Abernethy, B., Hanna, A., & Plooy, A. (2002). The attentional demands of preferred and non-preferred gait patterns. *Gait and Posture*, *15*, 256–265.

Adams, R., & Kleck, R. (2003). Perceived Gaze Direction and the Processing of Facial Displays of Emotion. *Psychological Science*, *14*(6), 644–647.

Adams, R., & Kleck, R. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, *5*(1), 3–11.

Adolphs, R. (2002). Recognizing Emotion From Facial Expressions: Psychological and Neurological Mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21–62.

Alferdinck, J., Hogervorst, M., Van Eijk, A., & Kusmierczyk, J. (2010). Mesopic vision and public lighting – A literature review and a face recognition experiment. *TNO Defence, Security and Safety*.

Archea, J. (1985). The Use of Architectural Props in The Conduct of Criminal Acts. *Journal of Architectural and Planning Research*, *2*(4), 245–259.

Baron-Cohen, S., & Cross, P. (1992). Reading the Eyes: Evidence for the Role of Perception in the Development of a Theory of Mind. *Mind & Language*, *7*(1–2), 172–186.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a “language of the eyes”? Evidence from normal adults, and adults with autism or Asperger Syndrome. *Visual Cognition*, *4*(3), 311–331.

Bensch, G., Peters, J., & Sievert, M. (2012). Fear of the Dark? – How Access to Electric Lighting Aff ects Security Attitudes and Nighttime Activities in Rural Senegal. *Ruhr Economic Papers*, *369*.

Bland, M. (2000). *An introduction to medical statistics* (3rd ed.). Oxford medical publications.

Blöbaum, A., & Hunecke, M. (2005). Perceived danger in urban public space: The impacts of physical features and personal factors. *Environment and Behavior*, *37*(4), 465–486.

Boomsma, C., & Steg, L. (2014). Feeling Safe in the Dark: Examining the Effect of Entrapment, Lighting Levels, and Gender on Feelings of Safety and Lighting Policy Acceptability. *Environment and Behavior*, *46*(2), 193–212.

Boyce, P. (2014). *Human factors in lighting* (3rd Editio). Crc Press.

Boyce, P. (2019). The benefits of light at night. *Building and Environment*, *151*, 356–367.

Boyce, P., Eklund, N., Hamilton, B., & Bruno, L. (2000). Perceptions of safety at night in different lighting conditions. *Lighting Research and Technology*, *32*(2), 79–91.

Boyce, P., & Rea, M. (1990). Security lighting: Effects of illuminance and light source on capabilities of guards and intruders. *Lighting Research and Technology*, *22*(2), 57–79.

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

BS 5489-1:2013. (2012). *Code of practice for the design of road lighting. Part 1: Lighting of roads and public amenity areas*.

BS 5489-1:2020. (2020). *Design of road lighting Part 1: Lighting of roads and public amenity areas — Code of practice*.

BS EN 13201-2:2015. (2016). *Road lighting Part 2: Performance requirements*.

Bull, P. (2016). *Posture & gesture* (Vol. 16). Elsevier.

Burton, A., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science*, *10*(3), 243–248.

Caminada, J. ., & Van Bommel, W. J. . (1984). New lighting criteria for residential areas. *International Lighting Review*, *13*(4), 350–358.

Chamilothori, K., Wienold, J., & Andersen, M. (2018). Adequacy of Immersive Virtual Reality for the Perception of Daylit Spaces: Comparison of Real and Virtual Environments. *LEUKOS*, 15(2–3), 203–226.

Chandra, S., & Bharti, A. (2013). Speed Distribution Curves for Pedestrians During Walking and Crossing. *Procedia - Social and Behavioral Sciences*, *104*, 660–667.

Chapman, A., Hawkins-Elder, H., & Susilo, T. (2018). How robust is familiar face recognition? A repeat detection study of more than 1000 faces. *Royal Society Open Science*, *5*(5).

Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, *112*(1), 155–159.

Commission Internationale de L'Eclairage. (2000). *CIE 136-2000 Guide to the Lighting of Urban Areas*.

Commission Internationale de L'Eclairage. (2010). *CIE 115:2010 Lighting of Roads for Motor and Pedestrian Traffic*.

Commission Internationale de L'Eclairage. (2016). *ILV: international lighting vocabulary 2nd Edition*.

Cook, M. (1971). *Interpersonal perception*. Penguin Books.

Cook, M. (1977). Gaze and Mutual Encounters Gaze in Social Encounter How long-and when-we look others "in the eye" is one of the main signals in nonverbal communication. *The Scientific Research Honor Society*, *65*(3), 328–333.

Crouzet, S., Kirchner, H., & Thorpe, S. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision*, *10*(4), 1–17.

Department of Transport. (2020). *Reported road casualties in Great Britain: 2019 annual report*.

Dias, C., Sarvi, M., Ejtemai, O., & Burd, M. (2015). Elevated desired speed and change in desired direction: Effects on collective pedestrian flow characteristics. *Transportation Research Record*, *2490*, 65–75.

Dong, M., Fotios, S., & Lin, Y. (2015). The influence of luminance, observation duration and procedure on the recognition of pedestrians faces. *Lighting Research and Technology*, *47*, 693–704.

Dunn-Rankin, P., Knezek, G., Wallace, S., & Zhang, S. (2014). *Scaling methods*. Psychology Press.

Ekkekakis, P., Hall, E., VanLanduyt, L., & Petruzzello, S. (2000). Walking in (affective) circles: Can short walks enhance affect? *Journal of Behavioral Medicine*, *23*(3), 245–275.

Ekman, P., & Friesen, W. (1967). Head and body cues in the judgment of emotion: a reformulation. *Perceptual and Motor Skills*, *24*(3), 711–724.

Ekman, P., & Friesen, W. (2003). *Unmasking the face: A guide to recognizing emotions from facial expressions*. Ishk.

Ellsworth, P., Carlsmith, J., & Henson, A. (1972). The stare as a stimulus to flight in human subjects: A series of field experiments. *Journal of Personality and Social Psychology*, *21*(3), 302–311.

Field, A. (2009). *Discovering statistics using SPSS* (Third Edit).

Fisher, B., & Nasar, J. (1992). Fear of Crime in Relation to Three Exterior Site Features: Prospect, Refuge, and Escape. *Environment and Behavior*, *24*(1), 35–65.

Fisher, B., & Sloan, J. (2003). Unraveling the fear of victimization among college women: Is the “shadow of sexual assault hypothesis” supported? *Justice Quarterly*, *20*(3), 633–659.

Fletcher-Watson, S., Findlay, J., Leekam, S., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception*, *37*, 571–583.

Foster, C., Hillsdon, M., & Thorogood, M. (2004). Environmental perceptions and walking in English adults. *Journal of Epidemiology and Community Health*, *58*(11), 924–928.

Fotios, S. (2019). Using Category Rating to Evaluate the Lit Environment: Is a Meaningful Opinion Captured? *LEUKOS*, *15*(2–3), 127–142.

Fotios, S., & Castleton, H. (2017). Lighting for cycling in the UK - A review. *Lighting Research and Technology*, *49*(3), 381–395.

Fotios, S., Castleton, H., Cheal, C., & Yang, B. (2017). Investigating the chromatic contribution to recognition of facial expression. *Lighting Research and Technology*, *49*, 243–258.

Fotios, S., Castleton, H., & Yang, B. (2018). Does expression choice affect the analysis of light spectrum and facial emotion recognition? *Lighting Research and Technology*, *50*(2), 294–302.

Fotios, S., & Gibbons, R. (2018). Road lighting research for drivers and pedestrians: The basis of luminance and illuminance recommendations. *Lighting Research and Technology*, *50*(1), 154–186.

Fotios, S., Hamoodh, K., & Clanton, N. (2019). What are you looking at? Testing nancy's rules for pedestrian interactions. *Proceedings of the 29th CIE Session*, 1669–1674.

Fotios, S., & Johansson, M. (2019). Appraising the intention of other people: Ecological validity and procedures for investigating effects of lighting for pedestrians. *Lighting Research and Technology*, *51*(1), 111–130.

Fotios, S., Mao, Y., Hamoodh, K., & Cheal, C. (2022). Using relative visual performance to predict performance of an interpersonal evaluation task with variation in adaptation luminance , observer age , skin tone , pavement reflection and interpersonal distance. *Lighting Research and Technology*, *0*, 1–12. https://doi.org/10.1177/14771535211069027

Fotios, S., Unwin, J., & Farrall, S. (2015). Road lighting and pedestrian reassurance after dark: A review. *Lighting Research and Technology*, *47*(4), 449–469.

Fotios, S., Uttley, J., & Fox, S. (2016). Exploring the nature of visual fixations on other pedestrians. *Lighting Research and Technology*, 1–11.

Fotios, S., Uttley, J., & Yang, B. (2015). Using eye-tracking to identify pedestrians' critical visual tasks. Part 2. Fixation on pedestrians. *Lighting Research and Technology*, *47*(2), 149–160.

Fotios, S., Yang, B., & Cheal, C. (2015). Effects of outdoor lighting on judgements of emotion and gaze direction. *Lighting Research and Technology*, *47*, 301–315.

Fotios, S., Yang, B., & Uttley, J. (2015). Observing other pedestrians: Investigating the typical distance and duration of fixation. *Lighting Research and Technology*, *47*(5), 548–564.

Fujikawa, T., & Kobayashi, K. (2012). The impact of background music and sound in decision making: A case study in experimental economics laboratory. *African Journal of Business Management*, *6*(7), 2449–2466.

Gao, X., & Maurer, D. (2011). A comparison of spatial frequency tuning for the recognition of facial identity and facial expressions in adults and children. *Vision Research*, *51*(5), 508–519.

Gifford, R. (1991). Mapping Nonverbal Behavior on the Interpersonal Circle. *Journal of Personality and Social Psychology*, *61*(2), 279–288.

Gover, A., Tomsich, E., Jennings, W., & Higgins, G. (2011). An exploratory study on perceptions of safety, fear of crime, and victimization experiences among faculty and staff at an urban university: A focus on gender. *Criminal Justice Studies*, *24*(1), 37–55.

Guyonneau, R., Kirchner, H., & Thorpe, S. (2006). Animals roll around the clock: The rotation invariance of ultrarapid visual processing. *Journal of Vision*, *6*(10), 1008–1017.

Hahn, C., O'Toole, A., & Phillips, P. (2016). Dissecting the time course of person recognition in natural viewing environments. *British Journal of Psychology*, *107*(1), 117–134.

Hall, E. (1966). *The hidden dimension* (Vol. 609). Anchor.

Hariyadi, A., & Fukuda, H. (2017). Study of Visibility Indices of Traditional Japanese Horizontal Blind "Sudare" Based on the Illuminance Different Level Using Physical and Digital Image Experiment. *Interior and Landscape Technology (BUILT)*, 9, 15–24.

Harpe, S. (2015). How to analyze Likert and other rating scale data. *Currents in Pharmacy Teaching and Learning*, *7*(6), 836–850.

Hausdorff, J., Yogev, G., Springer, S., Simon, E., & Giladi, N. (2005). Walking is more like catching than tapping: Gait in the elderly as a complex cognitive task. *Experimental Brain Research*, *164*(4), 541–548.

Hayhoe, M., Shrivastava, A., Mruczek, R., & Pelz, J. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *3*(1), 49–63.

HM Treasury. (2021). *Public Expenditure Statistical Analyses 2021*.

Hole, G., George, P., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, *31*(10), 1221–1240.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Ishihara, S. (1987). *The Series of plates designed as tests for colour-blindness*. Tokyo, Japan: Kanehara Shuppan Co., LTD.

Itier, R., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience and Biobehavioral Reviews*, *33*(6), 843–863.

Iwata, M., Ayama, M., Mori, T., Iwasaki, H., Kohko, S., Inoue, Y., Itsuki, H., & Kyoto, N. (2015). Appearance of human face and atmosphere of environment under LED street lights using different correlated colour temperature. *Proceedings of 28th Session of the CIE. Manchester, UK*, *28*, 1638–1646.

Johansson, M., & Rahm, J. (2015). Perceived lighting qualities and pedestrians' performance. *Proceedings of the 28th Session of the CIE, Manchester, United Kingdom*.

Johansson, M., Rosén, M., & Küller, R. (2001). Individual factors influencing the assessment of the outdoor lighting of an urban footpath. *Lighting Research and Technology*, *43*(1), 31–43.

Johnson, M., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19.

Johnson, M., & Mareschal, D. (2001). Cognitive and perceptual development during infancy. *Current Opinion in Neurobiology*, *11*(2), 213–218.

Johnston, A., Hill, H., & Carman, N. (1992). Recognising faces: Effects of lighting direction, inversion, and brightness reversal. *Perception*, *21*, 365–375.

Kawaguchi-Suzuki, M., Hogue, M. D., Khanfar, N. M., Lahoz, M. R., Law, M. G., Parekh, J., Zairina, E., Hong, J., Robles, Y. R., & Van Thang, V. (2019). Cultural sensitivity and global pharmacy engagement in Asia: India, Indonesia, Malaysia, Philippines, and Vietnam. *American Journal of Pharmaceutical Education*, 83(4), 651–664.

Kendo, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*, 22–63.

Kim, D. (2010). The Interactive Effects of Colors on Visual Attention and Working Memory: In Case of Images of Tourist Attractions. *International CHRIE Conference-Refereed Track*, 1.

Kirchner, H., & Thorpe, S. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776.

Knight, C. (2010). Field surveys of the effect of lamp spectrum on the perception of safety and comfort at night. *Lighting Research and Technology*, *42*(3), 313–329.

Knight, C., Van Kemenade, J., & Deveci, Z. (2007). Effect of outdoor lighting on perception and appreciation of end-users. *In 18th Biennial TRB Visibility Symposium*, 1–9.

Knoblauch, R., Pietrucha, M., & Nitzburg, M. (1996). Field studies of pedestrian walking speed and start-up time. *Transportation Research Record*, *1538*(1), 27–38.

Knutson, B. (1996). Facial Expressions of Emotion Interpersonal Trait Inferences. *Journal of Nonverbal Behavior*, *20*(3), 165–182.

Koskela, H. (1997). “Bold walk and breakings”: Women’s spatial confidence versus fear of violence. *Gender, Place and Culture: A Journal of Feminist Geography*, *4*(3), 301–320.

Kuo, H., Kuo, M., Tiong, I., Wu, P., Chen, Y., & Chen, C. (2011). Visual acuity as measured with Landolt C chart and Early Treatment of Diabetic Retinopathy Study (ETDRS) chart. *Graefe’s Archive for Clinical and Experimental Ophthalmology*, *249*, 601–605.

Kurien, D. (2010). Body language: silent communicator at the workplace. *IUP Journal of Soft Skills*, *4*.

Lagrange, R., & Ferraro, K. (1989). Assessing Age and Gender Differences in Perceived Risk and Fear of Crime. *Criminology*, *27*(4), 697–720.

Lebowitz, B. (1975). Age and Fearfulness: Personal and Situational Factors. *Journal of Gerontology*, *30*(6), 696–700.

Lee, G. R. (1982). Sex Diffrences in Fear of Crime Among Older People. *Research on aging, 4*(3), 284–298.

Li, T., & Yang, B. (2018). New Empirical Data for Pedestrian Lighting: Effect on Recognition Ability on Real 3D Facial Expression. *Conference: CIE 2018*, 106–111.

Lin, Y., & Fotios, S. (2015). Investigating methods for measuring face recognition under lamps of different spectral power distribution. *Lighting Research and Technology*, *47*, 221–235.

Loewen, L. J., Steel, G. D., & Suedfeld, P. (1993). Perceived safety from crime in the urban environment. *Journal of Environmental Psychology*, *13*(4), 323–331.

Macrae, C., Hood, B., Milne, A., Rowe, A., & Mason, M. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science*, *13*(5), 460–464.

Mantuano, A., Bernardi, S., & Rupi, F. (2017). Cyclist gaze behavior in urban space: An eye-tracking experiment on the bicycle network of Bologna. *Case Studies on Transport Policy*, *5*(2), 408–416.

Morsbach, H. (1973). Aspects of Nonverbal Communication in Japan. *Journal of Nervous and Mental Disease*, 157(4), 262–277.

Morton, J., & Johnson, M. (1991). Gesichtswahrnehmung bei Kindern_ Morton und Johnson.pdf. *Psychological Review*, *98*(2), 164–181.

Nasar, J., & Fisher, B. (1992). Design for vulnerability: Cues and reactions to fear of crime. *Sociology and Social Research*, *76*(2), 48–58.

Nasar, J., Fisher, B., & Grannis, M. (1993). Proximate physical cues to fear of crime. *Landscape and Urban Planning*, *26*(1–4), 161–178.

Nasar, J., & Jones, K. (1997). Landscapes of fear and stress. *Environment and Behavior*, *29*(3), 291–323.

Okud, S., & Satoh, R. (2000). Study on Establishing an Evaluation Method for the Visibility of the Human Face. *J. Illum. Engng. Inst. Jpn.*, *84*(11), 25–29.

Orne, M. (1962). On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications. *American Psychologist*, *17*(11), 776.

Oxley, J., Corben, B., Fildes, B., & Charlton, J. (2004). Older pedestrians: Meeting their safety and mobility needs. *In Proceedings of the Road Safety Research, Policing and Education Conference, Perth, Australia*, 14–16.

Pain, R. (2001). Gender, Race, Age and Fear in the City. *Urban Studies*, *38*(5–6), 899–913.

Peña-García, A., Hurtado, A., & Aguilar-Luzón, M. (2015). Impact of public lighting on pedestrians' perception of safety and well-being. *Safety Science*, *78*, 142–148.

Persike, M., Meinhardt-Injac, B., & Meinhardt, G. (2013). The preview benefit for familiar and unfamiliar faces. *Vision Research*, *87*, 1–9.

Plainis, S., Murray, I., & Charman, W. (2005). The Role of Retinal Adaptation in Night Driving. *Optometry and Vision Science*, *82*(8), 682–688.

Poulton, E. C. (1977). Quantitative subjective assessments are almost always biased, sometimes completely misleading. *British Journal of Psychology*, *68*, 409–425.

Rahm, J., & Johansson, M. (2018). Assessing the pedestrian response to urban outdoor lighting: A full-scale laboratory study. *PLoS ONE*, *13*(10), 1–20.

Raynham, P., & Saksvikronning, T. (2003). White Light and Facial Recognition. *The Lighting Journal*, *68*(1), 29–33.

Rea, M. (2000). Illuminating Engineering Society of North America. *The IESNA Lighting Handbook*.

Rea, M., Bullough, J., & Akashi, Y. (2009). Several views of metal halide and high-pressure sodium lighting for outdoor applications. *Lighting Research and Technology*, *41*(4), 297–320.

Rea, M., & Ouellette, M. (1991). Relative visual performance: A basis for application. *Lighting Research & Technology*, *23*(3), 135–144.

Remland, M. (2016). *Nonverbal communication in everyday life* (4th ed.). Sage Publications.

Rippe, J., Ward, A., Porcari, J., & Freedson, P. (1988). Walking for Health and Fitness. *JAMA: The Journal of the American Medical Association*, *259*(18), 2720–2724.

Roman, C., & Chalfin, A. (2008). Fear of Walking Outdoors. A Multilevel Ecologic Analysis of Crime and Disorder. *American Journal of Preventive Medicine*, *34*(4), 306–312.

Rombauts, P., Vandewyngaerde, H., & Maggetto, G. (1989). Minimum semi-cylindrical illuminance and modelling in residential area lighting. *Lighting Research and Technology*, *21*(2), 49–55.

Romnée, A., & Bodart, M. (2014). Does LED litghing improves pedestrians visual performances? *In Lighting Quality and Energy Efficiency*.

Rutishauser, U., Tudusciuc, O., Neumann, D., Mamelak, A., Heller, A., Ross, I., Philpott, L., Sutherling, W., & Adolphs, R. (2011). Single-unit responses selective for whole faces in the human amygdala. *Current Biology*, *21*(19), 1654–1660.

Schaller, M., Park, J., & Mueller, A. (2003). Fear of the Dark: Interactive Effects of Beliefs About Danger and Ambient Darkness on Ethnic Stereotypes. *PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN*, *29*(5), 637–649.

Smith, S., & Rea, M. (1980). Relationships between office task performance and ratings of feelings and task evaluations under different light sources and levels. *CIE Publication*, 50, 207–211.

Simons, R., Hargroves, R., Pollard, N., & Simpson, M. (1987). Lighting criteria for residential roads and areas. *CIE, Venice*, *53*, 35–39.

Sparrow, W. A., Bradshaw, E., Lamoureux, E., & Tirosh, O. (2002). Ageing effects on the attention demands of walking. *Human Movement Science*, *21*(5–6), 961–972.

Staats, H., & Hartig, T. (2004). Alone or with a friend: A social context for psychological restoration and environmental preferences. *Journal of Environmental Psychology*, *24*(2), 199–211.

Strick, M., Holland, R., & van Knippenberg, A. (2008). Seductive eyes: Attractiveness and direct gaze increase desire for associated objects. *Cognition*, *106*(3), 1487–1496.

Teufel, C., Westlake, M., Fletcher, P., & Hagen, E. (2019). A hierarchical model of social perception: Psychophysical evidence suggests late rather than early integration of visual information from facial expression and body posture. *Cognition*, *185*, 131–143.

Thoma, P., Bauser, D., & Suchan, B. (2013). BESST (Bochum Emotional Stimulus Set)- A pilot validation study of a stimulus set containing emotional bodies and faces from frontal and averted views. *Psychiatry Research*, *209*(1), 98–109.

Van Cauwenberg, J., Clarys, P., De Bourdeaudhuij, I., Van Holle, V., Verté, D., De Witte, N., De Donder, L., Buffel, T., Dury, S., & Deforche, B. (2012). Physical environmental factors related to walking and cycling in older adults: The Belgian aging studies. *BMC Public Health*, *12*(1), 142.

Van den Stock, J., & de Gelder, B. (2014). Face identity matching is influenced by emotions conveyed by face and body. *Frontiers in Human Neuroscience*, *8*(53), 1–8.

Vansteenkiste, P., Zeuwts, L., Cardon, G., Philippaerts, R., & Lenoir, M. (2014). The implications of low quality bicycle paths on gaze behavior of cyclists: A field test. *Transportation Research Part F: Traffic Psychology and Behaviour*, *23*, 81–87.

Vansteenkiste, P., Zeuwts, L., van Maarseveen, M., Cardon, G., Savelsbergh, G., & Lenoir, M. (2017). The implications of low quality bicycle paths on the gaze behaviour of young learner cyclists. *Transportation Research Part F: Traffic Psychology and Behaviour*, *48*, 52–60.

Veitch, J., Fotios, S., & Houser, K. (2019). Judging the Scientific Quality of Applied Lighting Research. *LEUKOS*, *15*(2–3), 97–114.

Vrij, A., & Winkel, F. (1991). Characteristics of the built environment and fear of crime: A research note on interventions in unsafe locations. *Deviant Behavior*, *12*(2), 203–215.

Ward, L. (1972). Category judgments of loudness in the absence of an experimenter-induced identification function: Sequential effects and power-function fit. *Journal of Experimental Psychology*, *94*(2), 179–184.

Wicherts, J., Veldkamp, C., Augusteijn, H., Bakker, M., van Aert, R., & van Assen, M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid P-hacking. *Frontiers in Psychology*, *7*(NOV), 1–12.

Willis, M., Palermo, R., & Burke, D. (2011a). Judging Approachability on the Face of It: The Influence of Face and Body Expressions on the Perception of Approachability. *Emotion*, *11*(3), 514–523.

Willis, M., Palermo, R., & Burke, D. (2011b). Social Judgments are Influenced By Both Facial Expression and Direction of Eye Gaze. *Social Cognition*, *29*(4), 415–429.

Wu, J., Zhang, Y., Sun, S., Li, Q., & Zhao, X. (2021). Generalized zero-shot emotion recognition from body gestures. *Applied Intelligence*, 1–19.

Yang, B., & Fotios, S. (2012). Inter-personal judgements for pedestrians at night: exploring information perceived at different distances. *Ingineria Iluminatului*, *1*(14), 31–44.

Yang, B., & Fotios, S. (2015). Lighting and recognition of emotion conveyed by facial expressions. *Lighting Research and Technology*, *47*(8), 964–975.

Yao, Q., Sun, Y., & Lin, Y. (2009). Research on facial recognition and color identification under CMH and HPS lamps for road lighting. *LEUKOS*, *6*(2), 169–178.

Yarbus, A. (1997). *Eye Movements and Vision*. Springer.

Yogev, G., Hausdorff, J., & Giladi, N. (2008). The role of executive function and attention in gait. *Movement Disorders*, *23*(3), 329–342.

Zhang, S., Liu, X., Yang, X., Shu, Y., Liu, N., Zhang, D., & Liu, Y. (2021). The Influence of Key Facial Features on Recognition of Emotion in Cartoon Faces. *Frontiers in Psychology*, *12*(August), 1–17.

Zhong, C., Bohns, V., & Gino, F. (2010). Good Lamps Are the Best Police: Darkness Increases Dishonesty and Self-Interested Behavior. *Psychological Science*, *21*(3), 311–314.