Routledge
Taylor & Francis Group

Volume 31 – Issue 1 – January 2023
Visual Cognition

Visual Cognition

Co-Editors: Benjamin Balas & Jonathan Flombaum

Routledge

ISSN: 1350-6285 (Print) 1464-0716 (Online) Journal homepage: www.tandfonline.com/journals/pvis20

Moving to maintain perceptual and social constancy

James T. Enns, Rachel C. Lin-Yang & Veronica Dudarev

To cite this article: James T. Enns, Rachel C. Lin-Yang & Veronica Dudarev (2023) Moving to maintain perceptual and social constancy, Visual Cognition, 31:1, 43-62, DOI: 10.1080/13506285.2023.2192992

To link to this article: https://doi.org/10.1080/13506285.2023.2192992

Published online: 30 Mar 2023.

Submit your article to this journal

Article views: 209

View related articles

View Crossmark data

Citing articles: 2 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=pvis20

VISUAL COGNITION
2023, VOL. 31, NO. 1, 43–62
https://doi.org/10.1080/13506285.2023.2192992

Routledge
Taylor & Francis Group

Check for updates

Moving to maintain perceptual and social constancy

James T. Enns, Rachel C. Lin-Yang and Veronica Dudarev

Department of Psychology, University of British Columbia, Vancouver, Canada

ABSTRACT

Past research on object constancy has tended to treat the viewer as a passive observer. Here we examine viewers' body and eye movements when they are asked to view photos of people in a gallery setting. Participants considered one individual in each photo, before indicating how socially connected they felt toward them and then moving to a spot in the gallery where they would be most comfortable when talking to them. Photographed individuals varied in their projected distance from the camera (near, far) and in their image resolution (sharp, slightly blurred). Results showed that participants looked more directly at near versus far individuals and at sharp versus blurred individuals. They also rated their social connection as stronger when the images were near versus far and sharp versus blurred. Where participants stood when making these ratings was strongly correlated with the projected distance of the images and with their ratings of social connection. These findings are discussed with regard to brain mechanisms for maintaining stability in our perceptions of geometric and social aspects of our world. They also highlight our inherent tendency to attribute qualities of our perceptual experiences to objects in that world.

ARTICLE HISTORY
Received 16 August 2022
Accepted 15 March 2023

KEYWORDS
Empathy; personal space; eye tracking; constancy; person perception

Humans are devoted to acting at least as much as they are to perceiving. This is now so well accepted that when a study begins with the assertion that action and perception are inextricably linked, it is easy to dismiss as a cliché. Yet a naïve reader might miss this point when surveying the literature on object constancy, that remarkable ability of the human visual system to perceive a stable world of objects. Objects are experienced, for the most part, as being of a constant size, shape, colour, and location, despite the brain receiving sensory information about those objects that are highly variable. Somehow the brain processes the fluctuating *media* (the sensory array) in a way that allows the consistent *message* (the perceived attributes) to be accurately attributed to specific objects in the environment. This is typically how studies of object constancy are framed for readers. The present study challenges this view.

The truism that action and perception are inseparably bound is easily missed because studies of object constancy rarely consider the actions humans take to maintain perception of a stable world. Consider the famed textbook examples of size constancy. When two equal-sized human paper figures are cut and pasted into two different locations of a tiled hallway depicted in depth, the figure toward the back of the hallway appears much larger than the figure near the front (Shepard, 1990). A stripped-down version of this effect occurs when two lines (presumptive railroad ties) are placed one above the other flanked by converging lines (presumptive rails); the line closer to the point of convergence appear to be longer (Gillam, 1973). The involuntary inferences our brain makes about the distance of the viewer from the depicted figures and from the converging lines are applied to the target objects. Notably, these demonstrations both involve pictures, and so no action on the part of the observer can influence what is seen. Moving closer to the picture, moving further away from it, or moving to one side does not change the relative relationships between the elements in the picture that lead to the interpretation of pictorial depth.

But not all demonstrations of size constancy rely on pictorial illusions. Here it is worth taking a short tour of neglected perception research. The Ames Room

CONTACT Veronica Dudarev vdudarev@mail.ubc.ca Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC, Canada V6T 1Z4

© 2023 Informa UK Limited, trading as Taylor & Francis Group

illusion, now found in many science museums as an interactive display, is a compelling demonstration of size constancy. Two real people of similar size, standing at the back of small room, appear very different in size when the interior of the room is viewed through a peephole. The peephole provides only a monocular view, making the situation resemble a picture in that respect, but without the benefit of other picture viewing cues indicating a flat surface. Unbeknownst to the viewer, the room is much longer on one side than the other, making one of the people more distant from the viewer. The viewer is fooled into thinking they are at equal distance because the angles cast on the viewer's retina by the walls, floor, and ceiling are rectilinear, capitalizing on a bias to experience rectangular edges over other arrangements in our familiar world. In reality, the room is an obtuse trapezoid when viewed from above. The requirement of peephole viewing severely limits the possible role of viewer action in experiencing the shape of the room. But not entirely. Ittelson and Kilpatrick (1951) wrote that the observer's incorrect assumptions about the room's shape could be overcome by using a long stick to actively probe various parts of the room while maintaining the peephole view. At first, the probing viewers were unsuccessful at touching points in the room they are aiming for. But after some practice, they were successful, not only touching target points with accuracy but seeing the room for its true shape. When that occurred, the two people also appeared to be equal in size. Note that throughout this gradual shift from illusion to reality, the pattern on the viewer's retina had not changed at all. But perception had. The observer's action contributed to a change in perception.

These and other demonstrations convinced Ittelson and Kilpatrick (1951, p. 55) that

> perception is never a sure thing, never an absolute revelation of 'what is.' Rather, what we see is a prediction of our own personal construction designed to give us the best possible bet for carrying out our *purposes in action* [added italics]. We make these bets on the basis of our past experience.

The authors argued that size and other constancies should not be defined with respect to the actual properties of the objects. They thought greater progress would be made by facing the conceptual challenge of understanding how *perceptual* constancy was achieved. They chose the *perceptual* over *object* constancy to emphasize that stable perception was an achievement of the observer, not a truth about the world. They added that "any behavior which tends to preserve the continuity and stability of the perceived world in the face of ever-changing relationships between observer and environment [can] be labeled 'constancy'" (Ittelson, 1951, p. 285).

Ittleson moved from this early work on size constancy to the broader field of environmental psychology. But the theme of the interrelatedness of perception and action remained strong. One oft-repeated line was that our " … environmental experience is that continuing product of an active endeavor by an individual to create for himself a situation within which he can optimally function and achieve his own particular pattern of satisfaction" (Ittelson et al., 1976, p. 187). These authors also referred to the viewer's "assumptive world" (pp. 291) because it is based on past experiences that give the viewer a "feeling of surety" that they can act effectively in the future.

More recent decades have seen a wealth of studies supporting Ittelson's claim that observer actions contribute to perception. Some studies showed that merely viewing a tool could activate the brain regions involved in acting on the tool (Ishibashi et al., 2016) and the behavioural motor plans for its use (Ellis & Tucker, 2000; Tucker & Ellis, 1998). Those motor activations in turn influence the attentional priority given to the parts of the object most closely associated with the actions (Borghi & Riggio, 2015; Bub et al., 2018). Another line of research showed that potential actions (i.e., affordances) affect our perception of the environment in profound ways. Bhalla and Proffitt's (1999) pioneering study showed that people wearing heavy backpacks, as well as those fatigued by previous exercise, or in poorer general health, judged the slant of a hill as steeper than people not wearing the backpacks, or not fatigued, or in better overall health. These findings spawned many other studies of the perception of environmental properties, including distance, size, and speed, and in tasks as wide-ranging as parkour, video gaming, tool-use, and virtual reality (MacIntyre et al., 2022; Proffitt, 2009).

These findings have not gone unchallenged. For instance, Durgin et al. (2009, 2012) showed that the backpack effect on slope perception was reduced when participants could attribute the reasons for

wearing them to causes unrelated to the slope estimates (see also Keric & Sebanz, 2021). Witt (2017) went on to use a clever virtual reality setup to show that the participant's agency (personal control) of their actions was critical in determining how much their actions influenced perception in an interactive video game. Witt's (2017) perspective is that information about action is integrated dynamically with perceptual information, with greater weight assigned to the action sources when the participant has maximum control over their action decisions.

Ittelson et al. (1976) noted, almost in passing, that if object and person perception follow the same general laws, social psychologists would be faced with an analogous problem of determining how human actions contribute to the establishment of a stable socio-emotional world. This line was picked up by Fritz Heider, who was working in roughly the same period as Ittelson (the middle of the twentieth century). In his PhD dissertation, Heider (1920) was already asking how it was possible for humans to perceive qualities of objects in the world, when all they had to work with were highly variable sensory signals in the mind. Not only that, but humans experience perceptual objects as being situated outside of themselves, in the world, not inside the brain where the sensory signals are registered. He concluded that human cognition must be based on attributions (blame assigned to causes), and that there was a strong human bias to attribute root causes to features of the external world rather than to features of our own biology.

Heider went on to note that inanimate physical objects (e.g., chairs, rocks) were given fundamentally different attributions by humans that animate objects (e.g., including other people). Persons, wrote Heider (1958, p. 21), "… are perceived as action centers and as such can do something to us. They can benefit or harm us intentionally, and we can benefit or harm them. Persons have abilities, wishes and sentiments; they can act purposefully, and can perceive or watch us." This makes animate objects unique from inanimate ones because we attribute agency to them. Moreover, our attribution of agency to other people includes the assumption (often unwarranted) that the expression of agency we see in them is rooted in their momentary attitudes and enduring personalities, and not in the environmental or situational factors they find themselves in. This idea later became widely known as the fundamental attribution error in person perception (Ross, 1977).

Heider (1958) also claimed that people use physical distance to convey social similarity, feeling socially closer to those who are more similar to them and vice versa. This perspective was later developed as Construal Level Theory (Trope, Liberman, & Wakslak, 2007; Liviatan, Trope, & Liberman, 2008; Trope & Liberman, 2010). One of its main claims is that a participant's close physical distance to objects is associated with more detailed, fine-grained, and contextualized representations of those objects. Conversely, farther physical distances are associated with coarser, more prototyped, and stereotypical representations. Levitian, Trope, and Liberman (2008) showed that social distance produces the same consequences as physical distance, with similar persons being considered on a more detailed, fine-grained level than people who are less similar to the observer. Won et al. (2018) demonstrated that participants' estimation of their spatial distance from another person is associated with perceived social distance. In most studies supporting this theory, participants were presented with written vignettes, pictures, or videos, and they responded by providing self-reported ratings of social connection (Trope & Liberman, 2010).

The current study

The present study is a proof-of-concept experiment inspired by Ittelson's (2007) challenge to measure the interactive effects of action and perception. It also follows Heider's (1958) advice that persons are distinct from inanimate objects in the attributions we ascribe to them. The study explores the actions that observers take to alter their body and eye position with respect to people they are viewing. Life-size images of actors were displayed on the wall of a room, at different depicted distances from the camera, and participants were asked to stand in the location they would choose to stand if they were interacting with the actors in real life. Thus, unlike previous studies supporting Construal Level Theory, where imagined distance is taken as a measure of social closeness after reading about a social situation (e.g., Duke & Nowicki, 1972; Martins et al., 2013; Sprecher & Fehr, 2005), here we used physical distance directly as a dependent variable. Participants were also asked to self-report on the degree of social

connection they felt with the actors so that we could assess the relations between the implicit measure of choosing where to stand and the explicit self-reports of social closeness.

Our study combined several technologies to study the action-perception interaction. First, we used digital displays of high-resolution, life-size, images of people, with the camera's depth of focus selectively manipulated to highlight one or another person in the photo. Second, we tracked the bodies of participants moving voluntarily to a preferred position in the room when viewing these images. Third, we used mobile eye tracking to monitor the gaze of participants throughout this experience. These technologies allowed us to create a micro-world that resembled familiar real-world activities for our participants in some respects and that differed from them in others. Viewing people at different apparent distances from oneself allowed the observer to approach or retreat, much as we do in many everyday situations when finding the comfortable distance between ourselves and others. However, the fact that observers viewed photos (and not live people nor virtual three-dimensional models) meant that apparent distance had to be inferred from all the visual cues available in a flat depiction (i.e., size, superposition, foreshortening, and image clarity), under the assumption that people are generally similar in size.

Participants were invited to participate in a gallery-viewing experience. We intended this experience to be similar to that of visiting an art gallery and freely interacting with the artefacts. Some objects prompt us to linger, some lead to closer inspection, others prompt us to look away and to move on. Figure 1 shows examples of the photographs viewed by the observers. In some respects, these photos resemble the human figures seen in the Ames Room, with equal-sized people subtending larger or smaller visual sizes, depending on their apparent distance from the viewer. However, they also differed from the Ames Room in that viewers could see many environmental cues consistent with their varying distance from the camera. There was no attempt on our part to provide the viewer with an illusion beyond that of the photograph itself.[1]

We instructed participants to imagine engaging with the people in the photos as though they were real, moving to a spot in the room where they would be most comfortable if they were actually talking to that person. This design allowed us to compare eye and body movements in response to the images at the same time that we made more traditional person perception measurements via the questions of social connection. It also gave participants considerable freedom in their choice of actions. In reporting the results, we will examine participants' behaviour in the rough order in which it

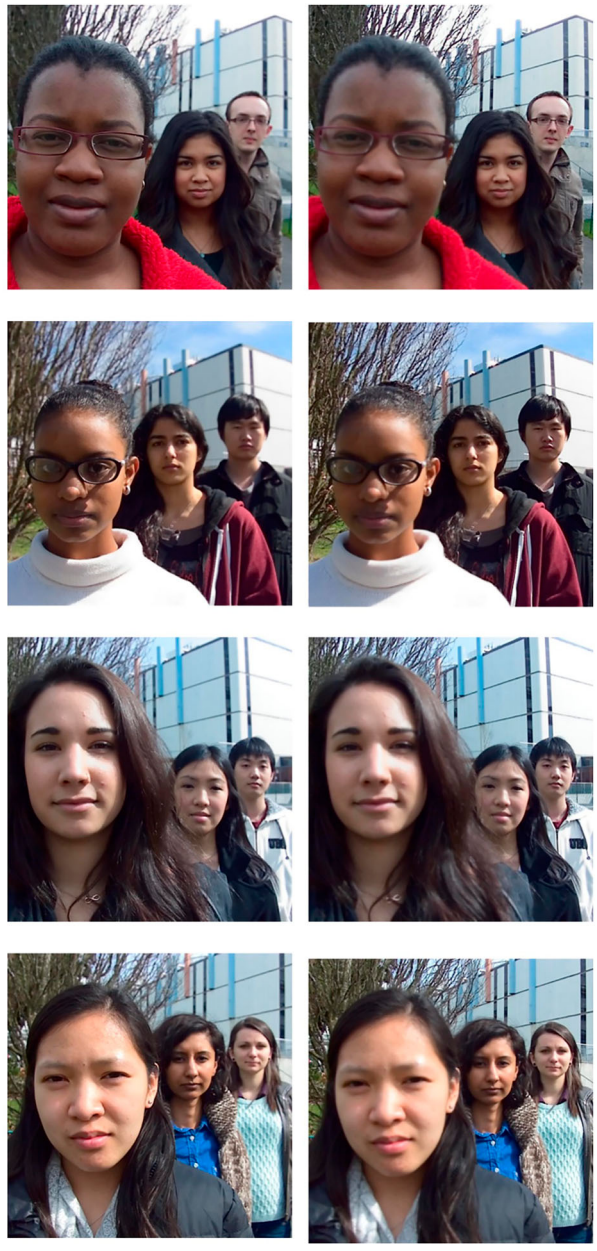

Figure 1. Examples of the photos used in the study. In the left photo in each row, the nearest person is at the camera's centre of focus; in the right photo, the farthest person is at the centre of focus. Across all photos, each of the models appeared equally often in each of the three positions and either in-focus (sharp) or out of focus (blurred).

unfolded. First, they looked at the target person in the photo in order to evaluate their possible social connection to that person. Then they moved to a location in the room that corresponded to their comfortable distance from that person. Finally, they verbally gave ratings on a 6-point scale to two questions assessing their social connection to the models.

Table 1 shows the questions, which were designed to index the perceived social connection participants felt with each actor using the concepts of empathy and altruism. "Empathic concern" is a spontaneous emotional response elicited by the perception of the expressed emotions of another in need (Batson, Lishner, & Stocks, 2015). Evidence shows that the strength of this response is directly related to the perceived similarity between an observer and the other (Perry, Bentin, Bartal, Lamm, & Decety, 2010). Empathic concern is thought to be the drive or motivation that results in altruistic behaviour (Batson et al., 2015). Studies in the comparative behaviour of other species suggest that this empathy-altruism connection is phylogenetically ancient (de Waal, 2008).

We organized our analyses of the eye tracking, rating, and body movement results to answer five questions. Our first hypothesis was that participants would look first, and then spend more time looking overall, at people depicted as physically closer to them than the same people depicted as more distant. Second, we expected that participants would look more at people rendered in sharp focus than the same people rendered slightly out of focus. Both of these outcomes would be consistent with what is known about how regions of conspicuity and clarity implicitly guide the viewer's gaze when examining portraits and other photos (DiPaola et al., 2010; Enns & MacDonald, 2013a; Enns & MacDonald, 2013b; Peterson et al., 2016; Reber et al., 1998).

A third hypothesis was that a consequence of increased looking influenced by relative distance and increased clarity would lead to a stronger perception of social connection (Enns & MacDonald, 2013b). This should be evident in the empathy and giving ratings participants gave to the models as a function of depicted distance and clarity. We did not have differential predictions for empathy ratings (the emotional response) versus giving ratings (the motivational drive) because we expected them to be closely related, and so we were prepared to treat any differences in the ratings as exploratory effects.

Our fourth hypothesis concerned body movement. We reasoned that participants might choose to stand further away from the screen when considering a near person in the photo and closer to the screen when they were considering a far person. This would be evidence that viewers act in the gallery setting to make their experience more similar to their daily experience interacting with people, where they move to maintain a constant comfortable personal distance (Hayduk, 1978; Sommer, 2002).

A fifth, and admittedly more exploratory hypothesis, was that participants would choose to stand closer to the screen when regarding models with whom they perceived a stronger social connection. Heider (1958) and Construal Level Theory reviewed earlier claim that the comfortable distance between oneself and other people is typically smaller for close acquaintances and larger for strangers with whom we feel less connected to (Hayduk, 1978; Sommer, 2002). It is also smaller for those we perceive to be more similar to us, even though they may be strangers (Perry et al., 2010), giving us a stronger experience of empathy and a greater motivational drive to help when in need. In the present study, all the models were unfamiliar to the participants and

Table 1. Participants provided a rating (6-point scale) on one of the empathy questions (randomly drawn) and one of the giving questions (randomly drawn) for each target model.

Empathy
P: This is someone I can relate to.
N: This person is not someone I can relate to.
P: I can tell how this person is feeling.
N: I have no idea how this person is feeling.
P: I would like to be this person's friend.
N: I would not like to be this person's friend.
P: It is easy to relate to this person.
N: It is hard to relate to this person.
P: I think we would get along.
N: I don't think we would get along.
P: This person would be easy to talk to.
N: This person would be hard to talk to.
Giving
P: I would give this person spare change if they were homeless.
N: I would walk the other way if this person was homeless.
P: I would donate to the World Wildlife Foundation if this person asked.
N: I would not donate to the World Wildlife Foundation if this person asked.
P: I would give to the Save the Children Fund if this person asked.
N: I would not give to the Save the Children Fund if this person asked.
P: I would sponsor this person in the Ride for the Cure (Cancer).
N: I would not sponsor this person in the Ride for the Cure (Cancer).
P: I would help this person change a flat tire.
N: I would not help this person change a flat tire.
P: I would donate blood if this person needed it.
N: I would not donate blood if this person needed it.

Half the questions in each set were worded positively (P), half were worded negatively (N).

so any findings for personal distance were attributed to perceived social connection. As with the ratings, we did not have differential predictions for whether the empathy or the giving ratings would be more strongly related to body movements, because of their close theoretical connections. As such, any differences found for body movement were treated as exploratory effects.

Method

Participants: photographic models

Twelve university students at the University of British Columbia volunteered in a half hour session to have their photos taken outdoors in groups of three. Each model signed a waiver giving permission for their images to be used in research on person perception and for their images to be published in describing that research. Their mean age was 21.4 years (SD = 3.6 years) and they self-identified as Black (2 female), East Asian (2 female, 2 male), South Asian (3 female, 1 male), and White (1 female, 1 male).

Participants: observers

Forty-two students at the University of British Columbia, with a similar age, gender, and ethnic composition as the models, volunteered through the Human Subject Pool at the University of British Columbia to participate in a half hour session for extra course credit. None of the observers recognized the models when they were asked during the study debriefing. Their mean age was 20.6 years (SD = 3.1 years) and they self-identified as Black (1 female), East Asian (12 female, 6 male), South Asian (6 female, 4 male), and White (9 female, 4 male). The only eligibility requirement was self-reported normal or corrected to normal vision. All observers read and signed a consent form prior to the study. All aspects of this study were approved in advance by the Behavioural Research Ethics Board at the University of British Columbia. The data were collected in 2015 as part of the second author's (R.C. Lin-Yang) undergraduate thesis project for her BSc degree. The total number of participants tested was therefore limited by the requirements to complete the project prior to the course deadline.

Stimulus materials

The photographs were taken with a Lytro camera (https://en.wikipedia.org/wiki/Lytro), which gives user's the opportunity to adjust the depth of focus after the photo has been taken. This allowed us to selectively highlight (with maximum photographic sharpness) either the person standing nearest to the camera or the person standing furthest away, without altering any other aspects of the photo. Focusing the image on the near person emphasized their prominence in the photo while further diminishing the appearance of the far person. Conversely, focusing the image on the far person increased their prominence while at the same time diminishing the prominence of the near person.

Examples of photos used in this study are shown in Figure 1. The 12 models were divided into four groups of three and each group was photographed in six different configurations, so that each individual appeared in the nearest, farthest, and middle positions. If we arbitrarily assign the labels 1, 2, 3 to the individuals in each group, and we use the label position to represent the spatial layout (nearest, middle, farthest), then the six configurations can be written as 123, 132, 213, 231, 312, 321. Notice that distance from the camera co-varied with a location in the photo, so that the nearest person was always standing on the left, and the farthest on the right. This generated 24 photos in all. The four examples in the top row all have the near person in sharp focus; the four examples in the bottom row have the far person in sharp focus.

The Lytro camera allowed us to selectively place each model at the centre of focus (to be the model with the sharpest image) or to have the same model be furthest from the centre of focus (to be blurred). We applied this selective sharpening or blurring only to the models in each photo standing nearest to the camera (on the left of each photo) or furthest from the camera (on the right). In total, this process generated 48 unique images for presentation in the study (4 sets × 6 configurations × 2 selective foci).

Each participant was asked to rate only one person in each photograph, and never the same person twice. This meant that each participant provided ratings for each of the 12 models, encouraging observers focus on the question "what do I think about this

unique individual, in comparison to the others I have seen and rated?", and not on the place a particular person was occupying in space. For each participant, we randomly selected images from the total set so that half the time the individual to be rated appeared near (on the left in the photograph), and half the time far (on the right). Crossed with this factor, half of the models to be rated were shown in sharp focus, and half the time blurred. This resulted in a 2 (position in the photograph: near vs. far) × 2 (focus: sharp vs. blurred) within-participant design, with 3 trials per participant in each of the four conditions.

The head of the near model (left) in each photo projected an image of 1.5 m on the screen (subtending 41 deg visual angle when the observer was standing on the spot marked X in the centre of the viewing gallery, 2 m from the screen). The head of the far model (right) projected an image of 0.5 m on the screen (subtending 14° visual angle at the same viewing distance). Standing at the velvet rope (.75 from the screen) made the model heads subtend from 70 (near) to 37 (far) degrees, while standing at the back of the room made the model heads subtend 16 (near) to 8 (far) degrees. We note for reference that a person with an average head size modelled as 23 cm in height subtends 14° of visual angle when viewed from 1 m (typical conversational distance) and 7° when viewed from 2 m (socially distanced).

The viewing gallery

The setting for the study was a large room, as illustrated in Figure 2 that was 3 m wide along the wall holding the projection screen, 4.5 m deep front-to-back, and 3 m high to the ceiling. Photos were projected with a Panasonic PT-DZ6700 projector with a resolution of 1920 × 1200, although the visible photos subtended a square aspect ratio of 1200 × 1200, which was 2 m on each side in the viewing gallery. The other three walls were surrounded by black curtains that ran from ceiling to floor. A velvet rope ran parallel to the projection wall 0.75 m from the screen to prevent observers from moving any closer to the screen. Behind the curtain on the left side was a table housing the desktop computer, a chair for the experimenter, and the various sensors that were placed on the observer before the study began. These included Dikablis mobile eye-tracking glasses (ergoneers.com/en/hardware/dikablis-glasses/) and a black baseball cap fitted with three infrared sensors in a triangle configuration for an Optitrak motion tracker (https://optitrack.com).

The Dikablis eye tracker has two cameras; one recording the wearer's right eye (with resolution of 384 × 288 pixels), and one facing forward to record the wearer's field of vision (with resolution of 786 × 576 pixels). Both cameras record at 25 frames per second. This means that the viewer's gaze location in the scene is recorded without being able to distinguish between the viewer's trunk, head, or eye position. The relatively slow frame rate also does not allow accurate differentiation between fixations (stable viewing locations) and saccades (rapid changes in location), meaning that each frame simply recorded the current eye position in the scene. A laptop computer connected to the desktop computer collected data from the Dikablis glasses via a long cable that ran down the observer's back to the floor and to the laptop. After the experiment, the video feeds were checked and corrected for calibration in Dikablis analysis software to ensure the accuracy of the eye tracking.

Presentation of the model photos was controlled with custom MATLAB software. The same computer sampled the motion of the observer's head using 12 infrared Optitrack cameras (1280 × 1024 resolution, 120 frames per second) and Tracking Tools software (optitrack.com/support/software/tracking-tools.html). The Optitrack system was calibrated at the start of each testing session to register a black cap with 3 reflective beads positioned in a triangle on the brim. The cameras began recording when the observer entered the curtained space entrance and ended when they left the space. Each presentation of a photo included a large arrow above the model to be rated on a given trial, as shown in Figure 2. This arrow randomly appeared above either the near (left) or the far model (right) in each photo.

Procedure

When observers arrived for the study, they were brought into the viewing gallery and given the consent form to read and sign. Observers were also asked to indicate their age, gender, and ethnic identity. They were then given verbal instructions for the rating task:

> We want to see how you feel about one person in each of the pictures you are about to see. Imagine that you were

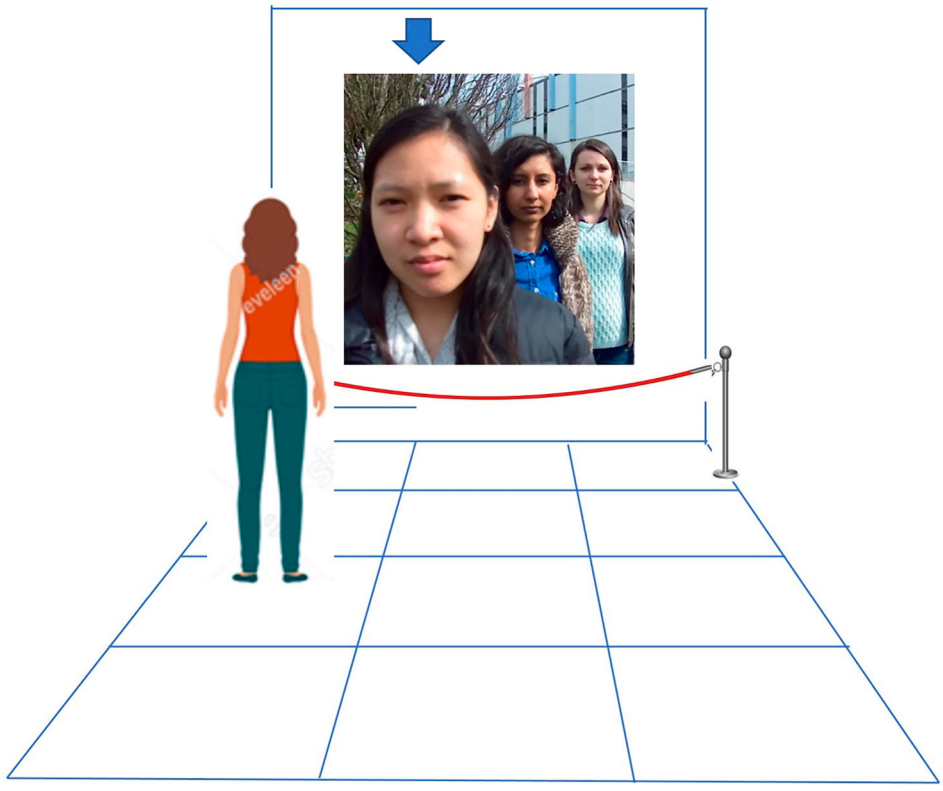

Figure 2. Illustration of the gallery setting for the study. The arrow above the screen indicates the target model. The figure facing the photo represents a study participant, who was asked to stand in the room as though they were engaging with the target in real life.

> given the opportunity to physically engage with each of these people you see on screen as if they were real. When considering how you feel about each person, please move to a spot in the room where you would be most comfortable if you were actually talking to them.

Once the instructions were understood the observer was fitted with the Dikablis glasses and the body tracking cap. The eye tracker was calibrated at the start of the experiment by having the observer stand on the X on the floor and then looking sequentially at four reference points that were projected near the corners of the photo location. The instructions were to fixate each point, starting with the upper left, and then to fixate each one in succession for 3 s, following a clockwise path. The body tracking cap was calibrated for each participant at the beginning of the experiment using a wand placed over the cap when the observer was standing on the spot marked with an X on the floor.

Room lights were off for the duration of the experiment, the only illumination coming from the projection screen, which allowed participants to see the whole room surrounded by black drapery and the screen separated from the viewer by the velvet rope. The observer began each trial standing on the spot marked on the floor with a taped X 2 m from the screen. They were reminded that

> when each picture comes on, feel free to move about the room so that you are best able to engage with this person as though they were real. Once you have moved to that location, and have spent 20–30 s viewing the person, please give a rating for each of two questions about this person. After you have finished the ratings, please walk back to the spot marked X on the floor and return your gaze to the plus sign at the center of the screen for 3 s.

The reference to the plus sign between trials allowed the experimenter to reposition the two images from the Dikablis eye-tracker so the pointer of the pupil was overlaid on the plus sign.

Each participant completed 12 trials in all, rating 12 unique models in all, equally divided between models standing near (left) or far (right) from the camera, and model images rendered in sharp or blurred focus. The two rating questions posed to participants for each target model were randomly drawn without replacement from the questions shown in Table 1. The empathy questions were equally divided between

positive and negatively worded questions, as were the questions concerning charitable giving. The negatively worded questions were reverse-coded for the purposes of data analysis (i.e., a negative item given a rating of 6 – strongly agree – was recoded as a 1).

Results

We obtained a complete set of data eye tracking and rating data from 38 of the 42 participants. Data from 4 participants were lost through calibration errors or a failure in coordinating the eye tracking data with the projected images. Data from an additional participant were lost through a calibration error in body tracking, leaving 37 participants for the body position data.

Eye tracking

Figure 3 shows the average horizontal eye position in the image over the first 20 gaze locations for all participants. This includes all gaze locations recorded during the first 10–15 s of viewing, including those recorded while looking around the room in order to decide where to stand before entering ratings. We examined the aggregated data over all 24 images and 38 participants, separated only by whether the target face was near (left side) or far (right side) and whether the target face was rendered in sharp detail (centre of focus) or was slightly blurred (outside the centre of focus).[2] We restricted our analysis to left-right dimension because the horizontal locations of the faces in the photos varied much more than their vertical locations.

Note first that the average gaze location began near the centre of the display on all trials. Yet, right from the outset, there was a bias to look more toward the near face on the left than the far face on the right, even on trials where the far face was targeted for ratings. This shows a bias to look initially at the larger near face on all trials, even when the arrow indicates the small far face is the target. Then, after 5–6 new gaze locations from the onset of a new display, the average gaze began to move more strongly in the direction of the target image: toward the left side of the image when the target was near and toward the right side of the image when the target was far. This directional tendency was stronger when the target image was depicted at the centre of the camera's depth of field (sharp) than when it was depicted outside of the camera's focus (blurred). This was especially noticeable on trials where the target face is in the far location. By the 12–13th gaze location, the strong directional tendency began to wane as participants allowed their eyes to examine other aspects of each image and began to answer the questions. This meant that participants spent more time overall looking toward the side of the near image than the far image. These two findings support our first hypothesis.

Our second hypothesis was confirmed for the smaller, further away images, but not for the larger, nearer images. An analysis of variance (ANOVA) examining the average gaze location in the range from 6 to 12 indicated an interaction between distance and clarity, $F(1, 36) = 26.35$, $p < .001$, partial $\eta^2 = .050$. For near targets, there was very little difference in looking behaviour for sharp and blurred models, $F < 1.0$. However, when far targets were rendered in sharp focus, they attracted gaze in their direction significantly more often than when far targets were outside the centre of focus (blurred), $F(1, 36) = 40.72$, $p < .001$.

Social connection: empathy and giving ratings

We first tested the empathy-altruism connection in the rating data by assessing their correlation. There was significant agreement among participants concerning the average empathy and giving ratings given to the 12 models, ICC2,38 = .637, $F(11) = 2.9$, $p = .001$ and ICC2,38 = .416, $F(11) = 1.9$, $p = .044$, respectively. The two subscripts on the intraclass correlations (ICCs) refer to the regression model and the number of raters (observers), respectively. Model 2 means that we used two-way mixed model as all targets were rated by all the 38 observers, who were randomly selected from a larger population of similar observers. As a group, participants rated some models significantly more highly than other models, when averaging over their depicted distances and clarity, $F_{empathy}(11, 492) = 3.22$, $p < .001$, partial $\eta^2 = .067$; F_{giving} $(11, 492) = 1.87$, $p < .04$, partial $\eta^2 = .040$. There was also significant agreement among the ratings given to the two questions across all models, participants, and conditions, $r(456) = .330$, $p < .001$. We also included the positivity of the wording as a

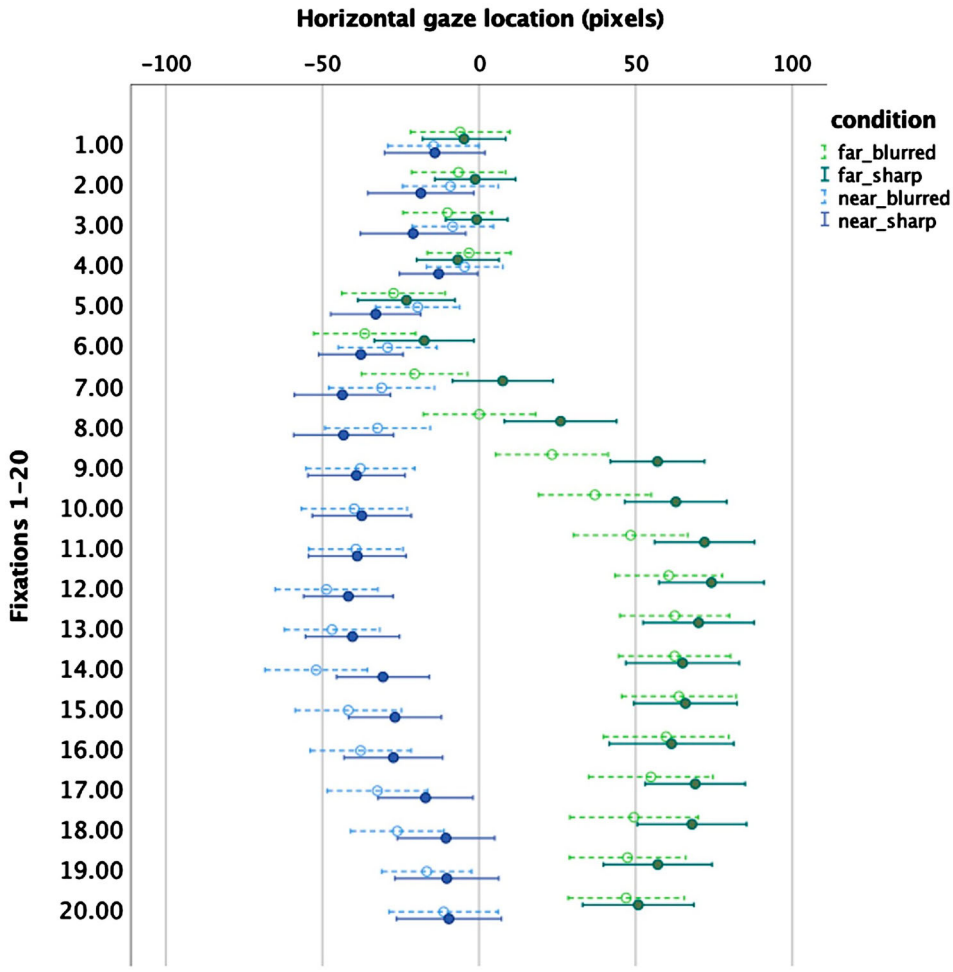

Figure 3. The average horizontal eye position in the image over the first 20 gaze locations for all participants. Error bars represent 95% confidence intervals about each mean.

factor in preliminary analyses and found no significant main effects or interactions. The Spearman rank order correlation between the mean empathy and mean giving ratings given to the 12 models was, $r(11) = .782$, $p > .001$. Thus, for the main analyses, we will report a combined measure (mean of the empathy and giving rating) to examine the relations between looking behaviour, ratings, and body movement. Secondary analyses examined the two ratings separately.

The average ratings given by participants to each of the target models were analysed with repeated-measures ANOVA, in which the fixed factors were apparent distance (near, far) and clarity (sharp, blur). The main measure for each participant in each of these four conditions was the average of three ratings given to the three models they saw in each of these conditions. Recall that across all participants, all models appeared in each of these conditions equally often and that each participant rated any given model only one time.

Figure 4(A) shows the average combined ratings as a function of the depicted distance and clarity of the target image. Near models were generally rated higher than far models, $F(1,37) = 12.91$, $p < .001$, partial $\eta^2 = .259$, and sharp models were generally rated higher than blurred models, $F(1,37) = 4.05$, $p < .05$, partial $\eta^2 = .099$. The interaction between these two factors was not significant, $F(1,37) = 0.05$, $p = .82$, partial $\eta^2 = .001$. Ratings for giving were generally higher than ratings for empathy, $F(1,37) = 41.28$, $p < .001$, partial $\eta^2 = .527$, and there was a significant 3-way interaction of rating type x distance x clarity, $F(1,37) = 9.23$, $p < .004$, partial $\eta^2 = .200$.

Figure 4(B,C) examined the 3-way interaction by looking separately at the empathy and giving ratings, respectively. For the empathy ratings, near models were generally rated higher than far models,

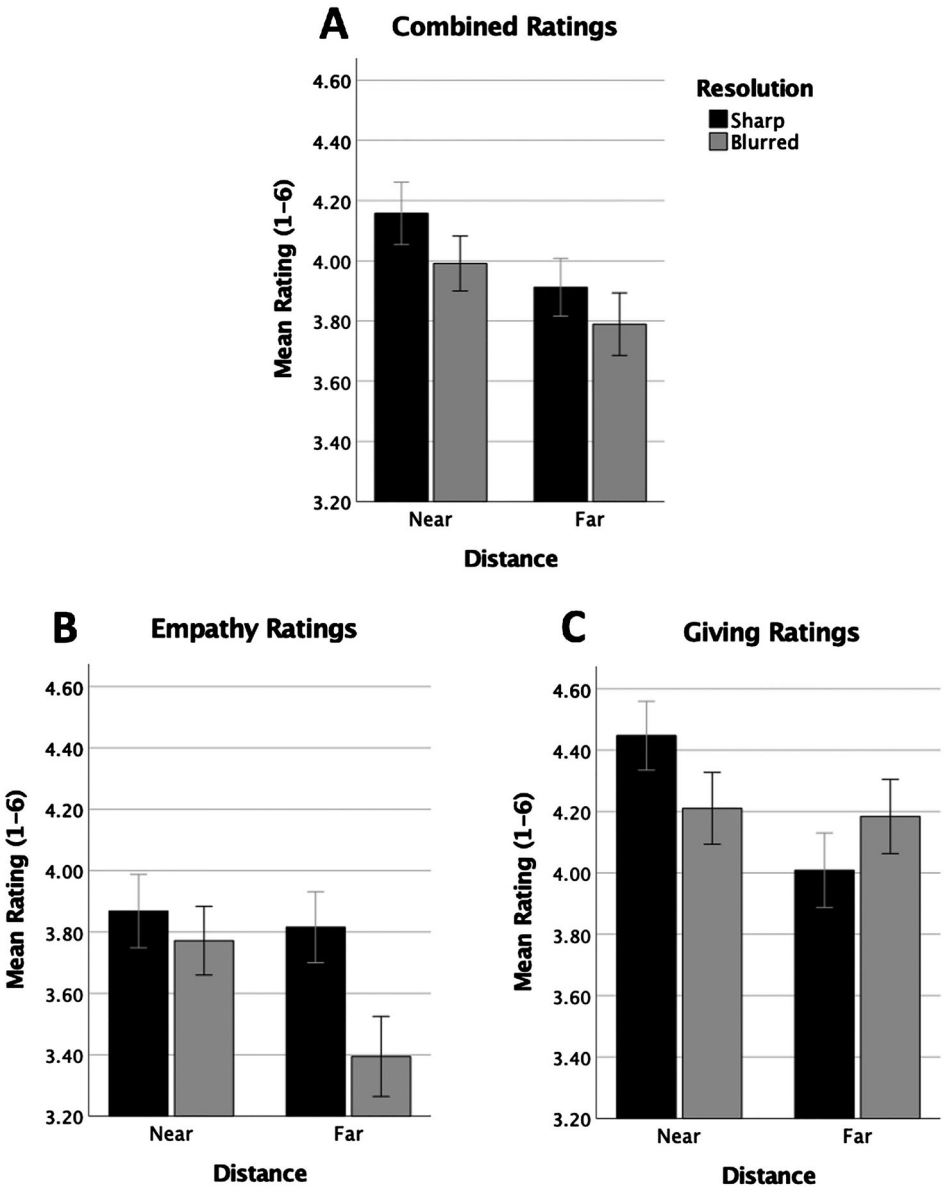

Figure 4. The top panel shows the mean combined ratings of social connection with the models as a function of depicted distance and image clarity. Error bars represent ±1 standard error above and below each mean. The data are shown separately for empathy ratings in the lower two panels.

$F(1,37) = 4.61$, $p < .04$, partial $\eta^2 = .111$, and sharp models were generally rated higher than blurred models, $F(1,37) = 5.77$, $p < .04$, partial $\eta^2 = .135$. A interaction between these two factors was not significant, $F(1,37) = 1.92$, $p = .18$, partial $\eta^2 = .049$. For the giving ratings, near models were also generally rated more highly than far models, $F(1,37) = 9.59$, $p < .01$, partial $\eta^2 = .206$. The main effect of clarity was not significant, $F(1,37) = 0.15$, $p = .70$, partial $\eta^2 = .004$, and the interaction between these two factors was not significant, $F(1,37) = 3.76$, $p = .06$, partial $\eta^2 = .092$.

In summary, empathy and giving ratings were higher when models were depicted as near than when the same models were depicted as further away. In addition, resolution of the models affected empathy ratings, such that when models were in sharp focus, they were rated more highly than when they were slightly blurred. The same effect of image clarity was not significant for giving ratings. Yet when the ratings were combined, the effects of image distance and clarity were both significant and additive with one another.

Body movement

Figure 5(A) shows the average floor location of the participants, shown separately for the near (blue) and far targets (green), and when the combined ratings of the target model (average of empathy and giving ratings) was a 2, 3, 4, 5, or 6. We combined the ratings of 1 and 2 because there were so few 1's in the data set (only 18 empathy and 3 giving ratings out of 444 ratings in total). It is clear that there are two factors exerting a strong influence on where participants chose to stand in order to rate the target models: participants stood nearer to the screen when the target was depicted as farther away and they stood nearer to the screen when their ratings indicated positive emotional connections to the target. These observations were supported by the following statistical analyses.

We first analysed the total distance participants moved from the start position by summing over the front-to-back (Y) and left-to-right (X) movements. Factorial ANOVA showed that the total distance participants moved was influenced by a significant main effect of depicted distance, $F(1, 424) = 122.14$, $p < .001$, partial $\eta^2 = .209$ and by a significant main effect of rating magnitude, $F(4, 424) = 6.14$, $p < .001$, partial $\eta^2 = .042$. Polynomial trend analyses showed the rating effect consisted of a

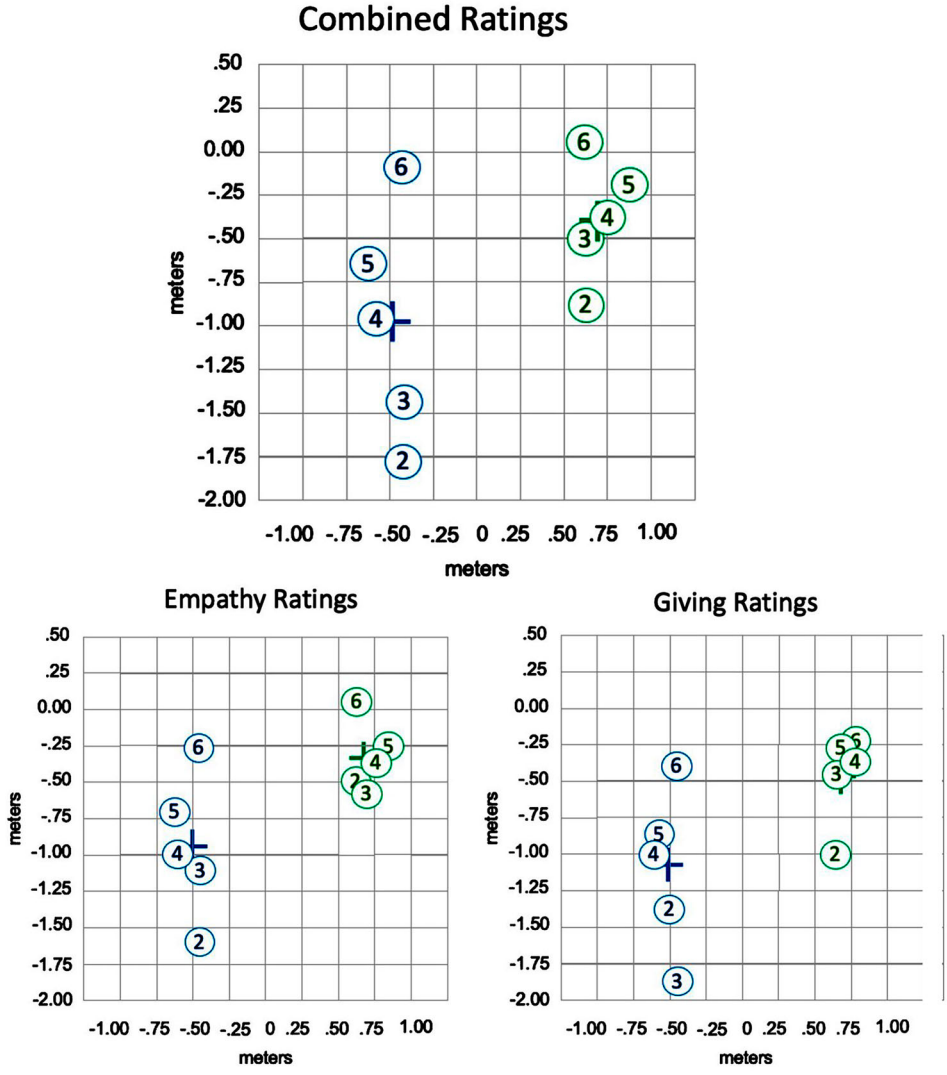

Figure 5. The top panel shows the mean floor location in the gallery when participants were rating the near model (blue circles) and the far model (green circles). The top of the chart represents the floor closest to the screen; 0, 0 marks the starting point for each participant. The number in each circle indicates the average floor location when participants rated the models with one of the five ratings (6 = strongest social connection, 2 = weakest social connection). The crosses indicate the mean across all ratings for near (blue) and far (green) models.

negative linear relationship between rating and distance, $F(1, 436) = 23.66$, $p < .001$ (no other polynomial contrasts were significant). The higher the rating, the closer the participant stood to the screen. There was no main effect of target clarity, $F = 1.09$, and there were no significant interactions among all these factors, all F-values < 1.5.

We next examined the position in the Y plane (distance from the screen) and X plane (horizontal position) separately. For the Y plane, there was a significant main effect of depicted distance, $F(1, 424) = 19.56$ $p < .001$, partial $\eta^2 = .040$, reflecting that participants stood closer to the screen for models depicted far from the screen than for models depicted near the screen. There was also a main effect of rating magnitude, $F(4, 424) = 8.94$, $p < .001$, partial $\eta^2 = .073$, indicating a negative correlation between distance to the screen and the mean rating. Position in the X plane on its own only revealed one significant main effect, with participants standing more on the left for near models and more to the right for far models, $F(1, 424) = 168.98$, $p < .001$, partial $\eta^2 = .270$, corresponding to the left and right image locations of the near and far models, respectively.

Figure 5(B,C) shows the same data separately for the two rating questions. The empathy ratings and the giving ratings were qualitatively similar to one another and therefore to the analyses reported above. The total distance score for empathy showed a significant main effect of depicted distance, $F(1, 424) = 124.18$, $p < .001$, partial $\eta^2 = .220$, and a significant main effect of rating magnitude, $F(4, 424) = 2.42$, $p < .05$, partial $\eta^2 = .017$, that contained a significant negative linear relationship between rating and distance, $F(1, 424) = 8.94$, $p < .001$ (no other contrasts were significant). There was no significant effect of target clarity and no significant interactions among these factors, all F-values < 1.0.

The analysis for giving ratings yielded similar results. There was a significant main effect of depicted distance, $F(1, 424) = 117.94$, $p < .001$, partial $\eta^2 = .208$, and of rating magnitude, $F(4, 424) = 4.07$, $p < .003$, partial $\eta^2 = .029$, that contained a negative linear relationship between rating and distance, $F(1, 424) = 11.09$, $p < .001$ (no other contrasts were significant). The main effect of target clarity and all interactions among these factors were not significant, all F-values < 1.0.

Discussion

The present study is a response to Ittelson's (2007) challenge to study perceptual constancy through the interactive measurement of action and perception. We observed that viewers moved their bodies to achieve stability in an imagined world, standing closer to images of people that were depicted as more physically distant and standing further away from people depicted to be nearer. The perception of apparent distance was supported in the photo by differences in image size, along with other pictorial depth cues such as foreshortening and partial occlusion. And the data showed not only that participants moved to maintain object constancy; they also moved to maintain social constancy. This aspect of the study was inspired by Heider's (1958) observation that maintaining perceptual constancy in a world where the objects are people, with agency, who can "do something to us," means that we will move in accordance with those perceptions. The results showed that participants reliably moved closer to the screen when giving an evaluation of someone they felt socially connected to, and they moved back from the screen when evaluating someone for whom they felt less empathy. This behaviour is completely consistent with a large body of research on the concept of personal space (Hayduk, 1978; Sommer, 2002).

These movement results were obtained in the context of participants being asked to consider how socially connected they felt toward a person depicted as standing in a small group with two other individuals. Unbeknownst to participants – because each participant only considered their social connectedness to 12 unique individuals – we examined the influence of projected scene distance and image clarity on the social connectedness ratings given by participants to these individuals. The three hypotheses we proposed at the outset of the study were each confirmed. First, we observed that participants looked first and then more often at individuals depicted as physically closer to them than the same individuals when depicted as more distant. Second, participants looked more directly at people rendered in sharp focus than the same people rendered slightly out of focus, especially when the image depicted them as further away. Both of these outcomes supported previous findings showing that local regions of selective conspicuity and clarity implicitly guide

the viewer's gaze when examining images (DiPaola et al., 2010; Enns & MacDonald, 2013a; 2013b; Peterson et al., 2016; Reber et al., 1998). The third hypothesis confirmed in the present data was that the increased looking that was caused by our manipulations of relative distance and clarity led directly to higher ratings of social connection (Enns & MacDonald, 2013b). In the section that follows we discuss the implications of these three findings, before returning to the results of our fourth and fifth hypotheses that concern body movement.

The role of image distance and clarity on person perception

Our interpretation of the finding that participants rated themselves as having greater empathy and more altruism for individuals when they were depicted as nearer (larger images), and when the images were rendered with greater clarity, is that participants' social perception was guided by implicit processes of inference that have been honed over a lifetime of experience.

Ordinarily, there is an inherent coupling between our selective attention and our personal interests because of the way our visual systems are designed. Every time we make an eye movement in the natural world, the image resolution of the scene on our retina changes dynamically. The high-resolution fovea of our eye, along with its surround by a lower-resolution visual periphery, guarantees that as we make eye movements, different regions of the scene become selectively sharp and other regions become blurred. This selective highlighting feature of natural vision is also supported by other design features. Consider lens accommodation. The lenses in our eyes are designed to change shape to assist in the formation of a sharp image on the fovea for the depth of field we are currently fixating (Campbell et al., 1958; Fisher & Ciuffreda, 1988). In addition to that, our two eyes work together to converge their gaze toward the same point in space. Image regions not at the point of converge are registered with lower resolution and so appear blurred. All of these factors combine to create a tight coupling in natural perception between image clarity and the viewer's spatial attention. Regions of the scene that are of greater interest to us are rendered with greater clarity on our eyes and in our minds. We see most clearly what we are most interested in.

Now consider how this natural interest-attention linkage is deployed when viewing pictures. Pictures are successful illusions (e.g., their content appears realistic to us) because they evoke many of the same perceptual processes that are evoked by a real three-dimensional scene. However, they also differ in important ways from real scenes. For one, we can experience their dual reality – we can see their flat surface at the same time that we experience their implied depth (Haber, 1980; Hochberg, 1980). Pictures have a sharply reduced dynamic range of luminance values when compared to their real-world equivalents – something we often don't notice until it is drawn to our attention (Hochberg, 1980). Most important for this study is that pictures are fixed in their image resolution values. This provides an opportunity for the natural interest-attention link to be reversed; for our selective spatial attention to be directed by image clarity rather than the other way around. By rendering some regions of a picture in greater detail through the depth of focus, the photographer and the artist implicitly invite the viewer to see the scene in a particular way. Regions rendered with greater clarity than other regions are being invited to be treated as more interesting.

What the present data show is that not only do participants look more often to people who are rendered to appear at a closer distance and with greater clarity, but they then take the next step, the attribution step of Heider (1958), to attribute their greater interest in these people to the characteristics of the people themselves. They come to see people rendered as larger and in sharper clarity as having dispositions more closely aligned with their own and as being more approachable and likable. This is evident in the higher ratings of social connection participants gave to models when they were rendered larger and with higher resolution images. It is important to be reminded again that these results were not specific to any particular model. Each model in the photos appeared equally often as a large (near) and as a small (far) image, and equally often as a sharp image and as a slightly out-of-focus image. Any given participant only saw a given model in one of these conditions, to preserve the uniqueness of the experience for them across the participants as a group. This experience included the illusion that

they were being asked to assess their social connection to real people and not being tested on factors of image quality.

These findings show that our perception of other people's social qualities in pictures is influenced by superficial image features in much the same way that it is influenced by a person's physiognomy (Willis & Todorov, 2006). As Heider (1920) observed long ago, humans are unable to escape the tendency to attribute qualities of their own perceptual experiences (including aspects of the media) to the meaningful content of those experiences. This is true in considering inanimate contexts as well (Nisbett & Wilson, 1977). Bar-Hillel (2015) demonstrated that the mere spatial position of consumer goods directs people to prioritize and praise qualities of random items in the particular spatial position in which they were displayed. This finding has been replicated numerous times and is now one of the guiding principles in marketing. In the present study case, the message concerns people with which we may potentially interact. In this case, the attributions we make about depicted people can be directly tied to our own actions, such as when we look longer, or more clearly, for reasons we are unaware of, and then attribute our behaviour to the person being viewed rather than to features of the media that delivered that message.

The outcome of person perception is mapped onto personal space.

The fourth and fifth hypotheses confirmed in this study concerned body movement. We observed that participants choose to stand further away from the screen when considering a near person in the photo and closer to the screen when they were considering a far person, consistent with their daily experience of maintaining a comfortable personal distance from people (Hayduk, 1978; Sommer, 2002).

Equally important, we saw that participants choose to stand closer to the screen when regarding the models they perceived to have a stronger social connection to. And as we have seen, this perceived social connection was influenced not only the idiosyncratic features of the depicted models, but also by their superficial image features of relative distance and clarity. These findings are strong support for Ittelson's (2007) idea that viewers act to be consistent with their "assumptive world," a world in which personal space – that comfortable distance between oneself and other people – is smaller for those we feel a stronger sense of social connection (Hayduk, 1978; Sommer, 2002).

Recent findings have also made it important to consider the possibility of demand characteristics playing a role in these data (Durgin et al., 2009; 2012). What may have been most obvious to our viewers is that we were interested in the social connection they imagined having with the depicted models. As such, the demand to approach or stay distant from the screen in order to signal perceived connection might have been obvious to them. What we think was much less obvious to them was that they should feel more connected to the same individual when the model was depicted with a larger image than when it was depicted with a smaller image. And even less obvious was that their connection should vary with the slight differences in depicted clarity, a factor that almost no one notices when viewing pictures such as these (DiPaola et al., 2010; Enns & MacDonald, 2013a; 2013b; Peterson et al., 2016). Future studies would do well to test these implicit effects directly by systematically controlling a participant's distance to far and near targets in order to see whether distance directly influences ratings when manipulated as an independent variable.

Recent neuroscience research has identified several brain regions that are critically involved in the establishment of personal space in primates (Graziano & Cooke, 2006; Kennedy et al., 2009). Especially relevant to the present findings are the ventral intraparietal (VIP) area and the precentral gyrus (PZ). Both of these regions have multimodal inputs and contain neurons with body-centred receptive fields that respond to objects that are near or looming toward the surface of the body. Electrical stimulation of these areas evokes defensive-like reactions and withdrawing movements in the monkey. Graziano and Cooke (2006) propose that these regions form the neural basis for personal space and the ethological phenomenon of a flight zone (Hediger, 1955). When the imaginary bubble of personal space is violated, a person steps away to reinstate the margin of safety. Numerous studies have shown that a person finding themselves in discomfort because of a potential threat or a lack of social connection will have an expanded personal space; a person in the friendly

company will have a reduced personal space (see Sommer, 2002).

A take-away message from the neurophysiological research of Graziano and Cooke (2006) is that brain areas that participate in the defense of the body surface are not restricted in their responses to variables such as perceived distance. Instead, they participate in highly complex and subtle behaviour. The same brain regions that govern body actions in response to physical looming are also involved in regulating behaviour that is informed by social and emotional information. This is entirely consistent with the results of the present study.

Another brain region that appears to play a strong role in the establishment of personal space is the amygdala (Kennedy et al., 2009). These authors reported a case study of an individual with complete amygdala lesions; one who lacked a consistent sense of personal space. In comparison, healthy individuals showed strong amygdala activation with increased personal proximity to a stranger. The authors proposed that the amygdala may be required to trigger the strong emotional reactions that occur when personal space is violated. Consistent with this interpretation, Gessaroli et al. (2013) reported a study of personal space in children diagnosed with autism spectrum disorder. These authors reasoned that if the amygdala plays a key role in underpinning personal space regulation (Kennedy et al., 2009), and if the amygdala is enlarged in children with autism Schumann et al. (2009), then this could contribute to the abnormalities of social fear and anxiety that are common in these children. The study used the popular stop-distance task, where participants are free to move to a comfortable position when interacting with an unfamiliar adult confederate. The results showed that children with the autism diagnosis were comfortable standing at a greater distance from the confederate when compared to typically developing children. Moreover, after interacting with the adult confederate to gain familiarity, the preferred personal distance was reduced for typically developing children, but the personal distance choices of the children with autism remained at the original larger distances (see also Mul et al., 2019).

Adding to the possible role of the amygdala in establishing personal space, Scheele et al. (2012) reported that injections of oxytocin, a neuropeptide known to reduce activity in the amygdala, led to a reduction in the preferred social distance between interacting partners in neurotypical study participants. Moreover, oxytocin injections also reduced the preferred social distance between interacting partners in individuals diagnosed with autism spectrum disorder (Andari et al., 2010).

It is important to point out that the arrows of influence between personal space and social comfort run in both directions. Not only does a reduced personal distance between people result in feelings of discomfort and the activation of defensive bodily reactions, but increasing the social connection between people leads to a reduced flight zone. Studies that measure the social-emotional constructs of compassion and intimacy rely very strongly on the metaphor of personal distance to index these constructs. For instance, in both the 10-item Compassion Scale (Martins et al., 2013) and in the 21-item precursor scale known as the Compassionate Love Scale (Sprecher & Fehr, 2005), the items that contribute most strongly to the first factor of compassion are "How much of your personal space would you share with a friend?" and "How much of your personal space would you share with a stranger that poses no threat to you?"

Also relevant to the present study are previously reported findings showing that perceived physical distance and one's feeling of social connection operate not only in live interactions but when viewing other people on television screens and in virtual reality situations (Bailenson et al., 2003; Llobera et al., 2010; Lombard, 1995; Probst, Rothe, & Hussmann, 2021; Wilcox et al., 2006). For example, Lombard (1995) asked participants to view news shows on televisions of various sizes and to stand at a "comfortable distance" to answer questions about how the shows made them feel. Study participants reported greater positive emotional responses to the people reporting the news on the screens, and to the environments depicted on the screens, when the screens were larger. Indeed, the participants preferred viewing position and emotional responses to the videos were influenced more strongly by the relative screen size of the images than by the participants' physical distance from the screen. Apparently, participants were taking the dual reality of television images into account in calibrating their behaviour with respect to the images. Much more important than their objective physical distance to the image (i.e.,

their recognition that they were viewing a flat image) was participants' perceptual distance from the people depicted in the images (i.e., their appreciation of the depicted image size relative to their own body size). This is consistent with the present findings. Studies of personal distance in virtual reality settings also show that violations of interpersonal space in these settings cause discomfort similar to that caused in natural settings. Not only are subjective ratings of discomfort influenced by life-sized stereoscopic images that appear too close (Probst et al., 2021; Bailenson et al., 2003), but physiological measures confirm a neurological basis for these experiences (Llobera et al., 2010; Wilcox et al., 2006).

Future directions

The present study clearly demonstrates that humans move their own bodies, when given the opportunity, in order to maintain perceptual constancy. Among the factors that influence these movements are both geometric (apparent distance) and social considerations (personal connectedness). Yet, the present study leaves many questions unanswered and raises entirely new questions. As such, we offer it primarily as a proof of concept that interactive measurements of perception and action are not only possible but important. Much more research will be needed to study these interactions in greater detail and in a wider range of contexts.

The present study tested only two different apparent distances and so cannot be used to address the question of how movement scales with changes in perceived distance. It was also limited to judgments made to photographs rather than to real people standing at various distances from the participant. All the models were strangers to the participants, and so the study says nothing about how social connectedness and personal distance interact when interacting with acquaintances and loved ones. The design also confounded apparent distance and screen location, as the nearer model was always on the left. Future studies should uncouple these two factors. Perhaps most importantly, the present study demonstrated that participants showed *sensitivity* to various visual factors in a social context (apparent distance, clarity), and in interaction with socially specific self-reports (empathy and charitableness), but the design did not offer a test of *specificity* for the social nature of the results. Perhaps these same results would be obtained if the objects being rated and interacted with were inanimate. We do not believe that is the case, but only future research with inanimate control objects will properly settle this question.

Despite these limitations, the findings prompt questions that may never have arisen, had we not designed a study in which human movement was one of the dependent variables. For example, there are obvious individual differences in the extent to which participants were influenced by apparent distance and clarity in their assessments of the models. Are these differences correlated with the degree to which their body movements reflected these assessments? Conversely, if participants are constrained in their movement, does their limited movement ability also become reflected in their perceptual judgments? Answering questions such as these will require exploring a larger portion of the parameter space, as well as testing a much larger number and greater diversity of study participants than we have at present. We hope the present study, with all its limitations, will inspire others to pursue these questions.

Notes

1. It is important to note that all pictorial art is an illusion (Hochberg, 1979; Sperandio & Chouinard, 2015). There is never any physical depth (variation in distance from the viewer) in a flat artwork or photograph; the interpretation we experience of looking into a three-dimensional world goes well beyond the physical data. That data itself conveys the duality of the experience; we simultaneously see the photograph as a flat image while also having a rich experience of the depicted depth relations in the image. Several researchers have noted that for this reason judgments of the real (retinal) size of an image are truncated when made in pictorial images versus when they are made from after-images, which do not share the dual reality of pictures (Gregory, 2008).
2. We were unable to perform a gaze analysis that was specific to each of the 24 different photos in the study because this was overwhelmingly difficult using the limited software that came bundled with the Dikablis eye tracker in 2015. Each photo had to be manually registered with the eye tracking data from a single trial and there were over 400 trials and each trial contained between 20 and 100 separate gaze locations. Instead, we analyzed the gaze data averaged over all 24 different photos and 38 participants, looking at the first 20 gaze locations recorded.

Acknowledgements

Author contributions to this paper included conceptualization, JTE; methodology, JTE; data collection, RL; statistical analysis, JTE, VD; writing, JTE, RL, VD; funding, JTE. Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Behavioral Research Ethics Board at the University of British Columbia (Approval LytroCam Study 2014). *Informed Consent Statement:* Informed consent was obtained from all participants involved in the study. Written informed consent was obtained from all 12 models photographed as the stimulus materials in this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by a Discovery Grant to JTE from the Natural Sciences and Engineering Research Council of Canada [grant number RPB-12R80005-2013].

Data availability statement

The data files supporting the results of this study can be found here: https://zenodo.org/badge/latestdoi/452038695

ORCID

Veronica Dudarev http://orcid.org/0000-0002-6383-8367

References

Andari, E., Duhamel, J. R., Zalla, T., Herbrecht, E., Leboyer, M., & Sirigu, A. (2010). Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. *Proceedings of the National Academy of Science, USA*, *107*(9), 4389–4394. https://doi.org/10.1073/pnas.0910249107

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, *29*(7), 819–833. https://doi.org/10.1177/0146167203029007002

Bar-Hillel, M. (2015). Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science*, *10*(4), 419–433. https://doi.org/10.1177/1745691615588092

Batson, C. D., Lishner, D. A., & Stocks, E. L. (2015). The empathy—altruism hypothesis. In D. A. Schroeder & W. G. Graziano (Eds.), *The Oxford handbook of prosocial behavior* (pp. 259–281). Oxford University Press.

Bhalla, M., & Proffitt, D. R. (1999). Visual-motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1076–1096. https://doi.org/10.1037/0096-1523.25.4.1076

Borghi, A. M., & Riggio, L. (2015). Stable and variable affordances are both automatic and flexible. *Frontiers in Human Neuroscience*, *9*, 351. https://doi.org/10.3389/fnhum.2015.00351

Bub, D. N., Masson, M. E. J., & Kumar, R. (2018). Time course of motor affordances evoked by pictured objects and words. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 53–68. https://doi.org/10.1037/xhp0000431

Campbell, F. W., Westheimer, G., & Robson, J. G. (1958). Significance of fluctuations of accommodation. *Journal of the Optical Society of America*, *48*(9), 669–669. https://doi.org/10.1364/JOSA.48.000669

de Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, *59*(1), 279–300. https://doi.org/10.1146/annurev.psych.59.103006.093625

DiPaola, S., Riebe, C., & Enns, J. T. (2010). Rembrandt's textural agency: A shared perspective in visual art and science. *Leonardo*, *43*(2), 145–151. https://doi.org/10.1162/leon.2010.43.2.145

Duke, M. P., & Nowicki, S. (1972). A new measure and social-learning model for interpersonal distance. *Journal of Experimental Research in Personality*, *6*(2-3), 119–132.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. (2009). Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin & Review*, *16*(5), 964–969. https://doi.org/10.3758/PBR.16.5.964

Durgin, F. H., Klein, B., Spiegel, A., Strawser, C. J., & Williams, M. (2012). The social psychology of perception experiments: Hills, backpacks, glucose, and the problem of generalizability. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1582–1595. https://doi.org/10.1037/a0027805

Ellis, R., & Tucker, M. (2000). Micro-affordance: The potentiation of components of action by seen objects. *British Journal of Psychology*, *91*(4), 451–471. https://doi.org/10.1348/000712600161934

Enns, J. T., & MacDonald, S. C. (2013a). The role of clarity and blur in guiding visual attention in photographs. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 568. https://doi.org/10.1037/a0029877

Enns, J. T., & MacDonald, S. C. (2013b). Photographic clarity and blur influences person perception. *Journal of Vision*, *13*(9), 421–421. https://doi.org/10.1167/13.9.421

Fisher, S. K., & Ciuffreda, K. J. (1988). Accommodation and apparent distance. *Perception*, *17*(5), 609–621. https://doi.org/10.1068/p170609

Gessaroli, E., Santelli, E., di Pellegrino, G., & Frassinetti, F. (2013). Personal space regulation in childhood autism spectrum disorders. *PLoS One*, *8*(9), e74959. https://doi.org/10.1371/journal.pone.0074959

Gillam, B. (1973). The nature of size scaling in the Ponzo and related illusions. *Perception & Psychophysics*, *14*(2), 353–357. https://doi.org/10.3758/BF03212404

Graziano, M. S., & Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia*, *44*(6), 845–859. https://doi.org/10.1016/j.neuropsychologia.2005.09.009

Gregory, R. (2008). Emmert's law and the moon illusion. *Spatial Vision*, *21*(3–5), 407–420. https://doi.org/10.1163/156856808784532509

Haber, R. N. (1980). How we perceive depth from flat pictures. *American Scientist*, *68*(4), 370–380.

Hayduk, L. A. (1978). Personal space: An evaluative and orienting overview. *Psychological Bulletin*, *85*(1), 117–134. https://doi.org/10.1037/0033-2909.85.1.117

Hediger, H. (1955). *Studies of the psychology and behavior of captive animals in zoos and circuses*. Criterion Books.

Heider, F. (1920). Zur Subjektivität der Sinnesqualitäten [On the subjectivity of sense qualities]. Unpublished doctoral dissertation. University of Graz, Austria.

Heider, F. (1958). *The psychology of interpersonal relations*. Wiley.

Hochberg, J. (1979). Some of the things that pictures are. In C. Nodine & D. Fisher (Eds.), *Views of pictorial representation: Making, perceiving and interpreting*. Praeger.

Hochberg, J. (1980). Pictorial functions and perceptual structures. In M. A. Hagen (Ed.), *The perception of pictures* (Vol. 2, pp. 47–93). Academic Press.

Ishibashi, R., Pobric, G., Saito, S., & Lambon Ralph, M. A. (2016). The neural network for tool-related cognition: An activation likelihood estimation meta-analysis of 70 neuroimaging contrasts. *Cognitive Neuropsychology*, *33* (3-4), 241–256. https://doi.org/10.1080/02643294.2016.1188798

Ittelson, W. H. (1951). The constancies in perceptual theory. *Psychological Review*, *58*(4), 285.

Ittelson, W. H. (2007). The perception of nonmaterial objects and events. *Leonardo*, *40*(3), 279–283. https://doi.org/10.1162/leon.2007.40.3.279

Ittelson, W. H., Franck, K. A., & O'Hanlon, T. J. (1976). The nature of environmental experience. In S. Wapner, S. B. Cohen, & B. Kaplan (Eds.), *Experiencing the environment* (pp. 187–206). Springer.

Ittelson, W. H., & Kilpatrick, F. P. (1951). Experiments in perception. *Scientific American*, *185*(2), 50–56. https://doi.org/10.1038/scientificamerican0851-50

Kennedy, D. P., Gläscher, J., Tyszka, J. M., & Adolphs, R. (2009). Personal space regulation by the human amygdala. *Nature Neuroscience*, *12*(10), 1226–1227. https://doi.org/10.1038/nn.2381

Keric, V., & Sebanz, N. (2021). Task construal influences estimations of the environment. *Frontiers in Human Neuroscience*, *292*(15). https://doi.org/10.3389/fnhum.2021.625193

Liviatan, I., Trope, Y., & Liberman, N. (2008). Interpersonal similarity as a social distance dimension: Implications for perception of others' actions. *Journal of experimental social psychology*, *44*(5), 1256–1269.

Llobera, J., Spanlang, B., Ruffini, G., & Slater, M. (2010). Proxemics with multiple dynamic characters in an immersive virtual environment. *ACM Transactions on Applied Perception*, *8*(1), 3:1–3:12. https://doi.org/10.1145/1857893.1857896

Lombard, M. (1995). Direct responses to people on the screen: Television and personal space. *Communication Research*, *22* (3), 288–324. https://doi.org/10.1177/009365095022003002

MacIntyre, E., Braithwaite, F. A., Mouatt, B., Wilson, D., & Stanton, T. R. (2022). Does who I am and what I feel determine what I see (or say)? A meta-analytic systematic review exploring the influence of real and perceived bodily state on spatial perception of the external environment. *Peer Journal*, *10*, e13383. https://doi.org/10.7717/peerj.13383

Martins, D., Nicholas, N. A., Shaheen, M., Jones, L., & Norris, K. (2013). The development and evaluation of a compassion scale. *Journal of Health Care for the Poor and Underserved*, *24*(3), 1235. https://doi.org/10.1353/hpu.2013.0148

Mul, C. L., Cardini, F., Stagg, S. D., Sadeghi Esfahlani, S., Kiourtsoglou, D., Cardellicchio, P., & Aspell, J. E. (2019). Altered bodily self-consciousness and peripersonal space in autism. *Autism*, *23*(8), 2055–2067. https://doi.org/10.1177/1362361319838950

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Perry, A., Bentin, S., Bartal, I. B. A., Lamm, C., & Decety, J. (2010). "Feeling" the pain of those who are different from us: Modulation of EEG in the mu/alpha range. *Cognitive, Affective, & Behavioral Neuroscience*, *10*, 493–504.

Peterson, J. J., Ringer, R. V., Riter, M., Sisco, E., De La Torre, M., Subedi, S., & Loschky, L. C. (2016). The effects of blur/clarity contrast on visual selective attention. *Journal of Vision*, *16*(12), 994–994. https://doi.org/10.1167/16.12.994

Probst, P. C., Rothe, S., & Hussmann, H. (2021, June). *Camera distances and shot sizes in cinematic virtual reality*. ACM International Conference on Interactive Media Experiences. https://doi.org/10.1145/3452918.3458804.

Proffitt, D. R. (2009). Affordances matter in geographical slant perception. *Psychonomic Bulletin & Review*, *16*(5), 970–972. https://doi.org/10.3758/PBR.16.5.970

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*(1), 45–48. https://doi.org/10.1111/1467-9280.00008

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, *10*, 173–220.

Scheele, D., Striepens, N., Güntürkün, O., Deutschländer, S., Maier, W., Kendrick, K. M., & Hurlemann, R. (2012). Oxytocin modulates social distance between males and females. *Journal of Neuroscience*, *32*(46), 16074–16079. https://doi.org/10.1523/JNEUROSCI.2755-12.2012

Schumann, C. M., Barnes, C. C., Lord, C., & Courchesne, E. (2009). Amygdala enlargement in toddlers with autism related to severity of social and communication impairments. *Biological Psychiatry*, *66*(10), 942–949. https://doi.org/10.1016/j.biopsych.2009.07.007

Shepard, R. N. (1990). *Mind sights: Original visual illusions, ambiguities, and other anomalies, with a commentary on the play of mind in perception and art*. WH Freeman.

Sommer, R. (2002). Personal space in a digital age. In R. B. Bechtel & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 647–660). Wiley.

Sperandio, I., & Chouinard, P. A. (2015). The mechanisms of size constancy. *Multisensory Research*, *28*(3-4), 253–283. https://doi.org/10.1163/22134808-00002483

Sprecher, S., & Fehr, B. (2005). Compassionate love scale for close others and humanity. *Journal of Social and Personal Relations*, *22*(5), 629–651. https://doi.org/10.1177/0265407505056439

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440.

Trope, Y., Liberman, N., & Wakslak, C. (2007). Construal levels and psychological distance: Effects on representation, prediction, evaluation, and behavior. *Journal of consumer psychology*, *17*(2), 83–95.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 830–846. https://doi.org/10.1037/0096-1523.24.3.830

Wilcox, L., Allison, R. S., Elfassy, S., & Grelik, C. (2006). Personal space in virtual reality. *ACM Transactions on Applied Perception*, *3*(4), 412–418. https://doi.org/10.1145/1190036.1190041

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Witt, J. K. (2017). A role for control in an action-specific effect on perception. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(10), 1791–1804. https://doi.org/10.1037/xhp0000447

Won, A. S., Shriram, K., & Tamir, D. I. (2018). Social distance increases perceived physical distance. *Social Psychological and Personality Science*, *9*(3), 372–380. https://doi.org/10.1177/1948550617707017