Fall 10-1-2021

From Eyes to Minds: Perceiving Perception, and Attending to Attention

Clara Colombatto
Yale University Graduate School of Arts and Sciences, clara.colombatto@yale.edu

Recommended Citation

Colombatto, Clara, "From Eyes to Minds: Perceiving Perception, and Attending to Attention" (2021). *Yale Graduate School of Arts and Sciences Dissertations*. 315.
https://elischolar.library.yale.edu/gsas_dissertations/315

Abstract

From Eyes to Minds: Perceiving Perception, and Attending to Attention

Clara Colombatto

2021

The most important visual stimuli that we encounter in everyday life may be other people, and in particular their eyes. We constantly monitor (and follow) where others are looking, and hundreds of studies have stressed the importance of eyes as uniquely powerful visual stimuli. This dissertation argues otherwise: The eyes are special only insofar as they signal deeper properties about the minds behind them—namely the nature and direction of others' attention and intentions. We empirically support this view in two ways: First, in studies of 'minds without eyes', we demonstrate how well-known gaze effects (such as prioritized processing of eye contact in the 'stare in the crowd') readily replicate without any eyes at all, when the direction of attention and intention is signified in other ways. Second, in studies of 'eyes without minds', we demonstrate that such gaze effects are reduced when the eyes do not signal any underlying pattern of attention and intentions, even though they clearly look like eyes, as in the phenomenon we have dubbed 'gaze deflection'. Finally, in a study of what we call 'unconscious pupillometry,' we also explore how the visual system automatically and unconsciously prioritizes others' degree of attention (vs. distraction). Ultimately, what matters is not just perceiving and attending to the relevant physical features, but rather perceiving perception, and attending to attention. Collectively, this work shows how seemingly reflexive visual processes can be surprisingly sophisticated, and how visual processing may extract not only physical attributes, but also mental states.

From Eyes to Minds:

Perceiving Perception, and Attending to Attention

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Clara Colombatto

Dissertation Director: Brian J. Scholl

December 2021

Contents

Author List

Benjamin van Buren and Brian J. Scholl contributed to Chapter 2, which was published as:

Colombatto, C., van Buren, B., & Scholl, B. J. (2019). Intentionally distracting: Working memory is disrupted by the perception of other agents attending to you—even without eye-gaze cues. *Psychonomic Bulletin & Review, 26,* 951–957.

Benjamin van Buren and Brian J. Scholl contributed to Chapter 3, which was published as:

Colombatto, C., van Buren, B., & Scholl, B. J. (2020). Gazing without eyes: A 'stare-in-the-crowd' effect induced by simple geometric shapes. *Perception, 49,* 782–792.

Yi-Chia Chen and Brian J. Scholl contributed to Chapter 4, which was published as:

Colombatto, C., Chen, Y. -C., & Scholl, B. J. (2020). Gaze deflection reveals how gaze cueing is tuned to extract the mind behind the eyes. *Proceedings of the National Academy of Sciences, 117,* 19825–19829.

Brian J. Scholl contributed to Chapter 5, which is in press as:

Colombatto, C., & Scholl, B. J. (in press). Unconscious pupillometry: An effect of 'attentional contagion' in the absence of visual awareness. *Journal of Experimental Psychology: General.*

Listing of figures

Acknowledgments

This work would have not been possible without the many mentors, colleagues, friends, and family members who have supported me throughout my time in graduate school.

First and foremost, I am immensely grateful to my advisor Brian Scholl, for the endless attention and countless hours he devoted to the development of this work and to my growth as a scientist. Brian's thoughtfulness, intellectual taste, scientific rigor, expertise, and eloquence constitute the foundations of this work, and there is perhaps no gift more valuable or teaching more inspiring than the extraordinary care he dedicates to mentoring his students and to unraveling the mysteries of the mind. I still cannot believe I get to be your student.

I am also profoundly thankful to Molly Crockett for her invaluable guidance, fierce strength, and relentless support; to Ilker Yildirim, for his contagious creativity and pure love for science; to Steve Chang and Nick Turk-Browne, for their critical insights, which contributed greatly both to the current work and to my path ahead; to the other faculty members who have contributed to my education over the years—including John Bargh, Ty Cannon, BJ Casey, Joshua Knobe, L.A. Paul, and Dave Rand; and to the mentors who first showed me the beauty of cognitive science—including Greg Appelbaum, Kait Clark, Scott Huettel, Steve Mitroff, Karen Neander, and Walter Sinnott-Armstrong.

I feel fortunate to have been surrounded by many brilliant colleagues who make Yale Psychology an incredibly vibrant place to work and grow. I am especially grateful to the many past and current members of the Yale Perception & Cognition Lab, and especially to

Yi-Chia Chen, for being a precise and passionate scientist, and a patient and generous teacher; to Stefan Uddenberg and Benjamin Van Buren, for being supportive collaborators and loyal friends; to Chaz Firestone, Hannah Raila, Monica Rosenberg, and Emily Ward, for being the role models I am lucky to call friends; and to Conor Downey, Rui Zhe Goh, and Joanna Zukin, for watching you grow has been one of the greatest joys of graduate school. I also thank the members of the Cognitive & Neural Computation Lab, the Crockett Lab, the Human Neuroscience Lab, the Turk-Browne Lab, as well as the Department staff—and particularly Lynn Butler, Andrea Chamba, Patty McClure, and Lauretta Olivi.

Thank you to the friends who made New Haven home—including Dorsa Amir, Kevin Anderson, Matthew Archer, Kendall Arslanian, April Bailey, Adam Chekroud, Julia Clancy, Meghan Collins, Cameron Ellis, Asu Erden, Michael Farruggia, Danielle Gerhard, Kat Graves, Sam Haller and the Wednesday crew, Claire Haugh, Matt Jordan, Sharif Kronemer, Tess Lanzarotta, Qi Lin, Francesca Mandino, Julia Marshall, Thomas O'Connell, Luca Perdoni, Teresa Rossi, Brynn Sherman, Jen Siegel, Jeff Wammes, and Sami Yousif.

Finally, I am eternally grateful to the Lawrences—Abi, Charlotte, Henry, and Lisa; to my friends on the other side of the ocean—especially Anna, Filippo, Francesco, Matteo, Rossana, Viola, Violetta, Virginia, Tomà; and to my family—my aunt Sarah, my brothers Alberto and Federico, and my mother Elisa—for their unwavering support and unconditional love.

1

Introduction

The most powerful visual stimuli we encounter in our everyday lives are arguably other people: Identifying agents and tracking their behavior is of obvious importance to our fitness. But not all parts of other people are equally important: Some body parts like legs or elbows seem of lower biological significance compared to highly informative ones like people's faces—and in particular their *eyes*. Accordingly, a vast literature in psychology has explored how our minds seem especially sensitive to others' eyes: Compared to other stimuli, both social and nonsocial, eyes are detected faster, hold our attention for longer, and have many influences on the way we perceive and remember other people. As a result, hundreds of papers in social perception all converge on the conclusion that the eyes are the social signal *par excellence*, unparalleled in terms of both their salience and their impact on our mental lives.

This thesis argues otherwise: The eyes are not so special after all, and they are special only insofar as they signal deeper properties about the *minds* behind them—namely the direction of others' attention and intentions. Across a wide range of perceptual processes, from unconscious processing to impressions of temporal order, the work presented here demonstrates that the effects of eye gaze are driven not so much by the superficial appearance of the eyes, but rather by patterns of attention and intentions that they (only sometimes) signify. This conclusion may seem surprising, since perception is typically thought to extract properties like color and shape, while mental states are typically thought to result from high-level judgments and reasoning—a divide that is also reflected in the sociology of our field, since the study of mental state inferences has largely been the province of social psychology rather than psychophysics. But the work presented here suggests that perception itself may traffic in the perception of others' perceptual and attentional states. In this way, our visual experience seems to be driven not so much by what other agents look like superficially, but rather by deeper properties of their minds, such as their perceptual and attentional states.

The contribution of this research program is twofold. First, the possibility that perception might extract others' mental states has several implications for our understanding of perception, since visual processing can extract sophisticated models of others' minds even in the seemingly rapid and reflexive tasks explored in this work. And conversely, the suggestion that our apprehension of others' mental states is rooted in perception implies that mental states are not just something that we can think and reason about, but also something that we can *see*, and that might be truly integrated into how we experience the world in the first place. But even beyond its implications for the nature of other cognitive processes, this work begins to uncover our ability to apprehend others' perceptual and attentional states—properties that seem foundational not just to our own visual experience as demonstrated here, but also to many other aspects of our social lives, from conversations to legal and moral judgments. In fact, this form of social perception bridging automatic perceptual processes and sophisticated social impressions raises the intriguing possibility that by perceiving itself, perception may serve as a scaffold for other mentalizing abilities, insofar as the realization that others can see may function as a precursor for understanding their minds.

In the remainder of this introduction, I will first provide an overview of the numerous ways in which the eyes seem special, with a particular focus on the effect of others' gaze on automatic mechanisms of perception, attention, and memory. I will then ask: Why might the eyes be special? While vision science has typically taken the effects of gaze to be driven by the eyes being especially visually salient stimuli, I will propose a new theory wherein these effects are really driven by the perception of the attention and intentions of the *minds* behind the gaze. Finally, I outline two key predictions of our framework, and introduce the logic of the experiments presented in the remaining chapters.

1.1 The Most Important Visual Stimulus?

Of the many stimuli we encounter in everyday life, perhaps the most important are others' faces: If much of our waking lives is spent in the presence of other people (U.S. Bureau of Labor Statistics, 2019) or thinking about others (Mar, Mason, & Litvack, 2012; Song & Wang, 2012), then faces, as the proximal means by which we identify, observe, and remember other agents, seem crucial. And indeed, we learn a lot about others from just glancing at their faces, from demographic characteristics such as age or gender, to more subjective impressions of familiarity and emotional expressions, and even more ineffable impressions of personality traits and mental states. Faces are of course complex stimuli composed of multiple internal features (e.g., nose, mouth, ears, hair, ...), and while face perception seems to rely on global processing of local features as a configural unit (Yin, 1969; for a review, see Maurer, Grand, &

Mondloch, 2002), one feature in particular seems most important: the eyes.

Despite their relatively small size compared to other facial features, the eyes seem especially salient to us: Eye-tracking studies have revealed that when viewing others' faces we tend to mostly focus on the eye region (Buswell, 1935; Yarbus, 1967), such that up to 80 % of fixations fall within the eye region only (Williams & Henderson, 2007). This tendency appears to be automatic, insofar as it occurs regardless of the task at hand (Davies & Hoffman, 2002; Henderson, Williams, & Falk, 2005; Janik et al., 1978; Walker-Smith, Gale, & Findlay, 1977), or the complexity of the scenes we are looking at (Birmingham, Bischof, & Kingstone, 2008). Both developmental and comparative work have suggested that this looking preference might be especially heightened in humans. First, the eyes seem special from the very beginning of life: Newborns prefer to look not just at faces (Goren, Sarty, & Wu, 1975; Johnson, 2005; Johnson et al., 1991), but specifically at the eye region (Batki et al., 2000; Farroni et al., 2002, 2005), starting as early as 7 weeks of age (Haith, Bergman, & Moore, 1977). In contrast, this preference for the eye region seems reduced if not absent in other species, including other great apes such as gorillas and orangutans (Kano, Call, & Tomonaga, 2012), and even our closest relatives—chimpanzees (Kano & Tomonaga, 2010).

If faces are so informative, most properties we can 'read' about others reside in fact in the eyes. For example, biometrical studies have revealed that the facial region that varies the most across individuals is indeed the eye region, including eye color, eye shape, inter-ocular distance, inter-pupillar distance, and others (Farkas, 1994). This significant inter-subject variability in morphology makes the eyes an exceptionally reliable cue for identity recognition: Performance in facial recognition tasks is significantly impaired when the eye region is occluded, more than when other facial features such as the nose or mouth are (Caldara et al., 2005; McKelvie, 1976; Schyns, Bonnar, & Gosselin, 2002; Sekuler et al., 2004). Similarly, accuracy on facial recognition tasks decreases when participants are forced to fixate away from the eyes (Peterson & Eckstein, 2012). And perhaps even more strikingly, the eyes seem crucial for detecting the presence of others in the first place—as demonstrated by face detection experiments wherein masking the eyes, but not other local features such as the mouth, results in significant performance impairments (Lewis & Edmonds, 2003). The eye region is not just helpful for identifying others, but also for extracting specific attributes, from gender (Gosselin & Schyns, 2001; Vinette, Gosselin, & Schyns, 2004) to disease immunity, which varies with the brightness of the sclera (Provine, Cabrera, & Nave-Blodgett, 2013; Russell et al., 2014), and the width of the dark limbal ring separating the sclera from the iris (Peshek et al., 2011). That the eyes are key determinants for impressions of age and health is indeed reflected in the massive market for cosmetic products for the eye region such as dark eyeliners and mascaras that enhance its apparent contrast, or eye shadows that enlarge its apparent size.

And even beyond such relatively stable traits, the eyes also convey information about more transient states such as emotions (Wagenbreth et al., 2014). The detailed musculature of the eyes and eyebrows is a major contributor to all primate facial expressions (Ekman & Friesen, 1971; van Hooff, 1967; for a review, see Itier & Batty, 2009), e.g., the wide eyes characteristic of fearful and surprised faces (Adolphs et al., 2005), the squinted eyes of angry and disgusted faces (Calder et al., 2000; Chapman & Anderson, 2012; Smith et al., 2005), the lowered gaze of sad faces (Semyonov et al., 2019), the wrinkled eye corners of genuinely joyful faces (Duchenne, 1990; Ekman & Friesen, 1982), and more complex configurations of sophisticated expressions such as jealousy, guilt, hate, suspicion, awe, and cowardice (Baron-Cohen et al., 1997, 2001; Lee & Anderson, 2017).

1.2 Gazing Direction

But perhaps the most obvious way in which others' eye gaze is informative is that it signals where in the environment others are looking. Our ability to discriminate others' looking direction is truly remarkable: We are able to detect averted gazes with eye rotations that are as small as 0.75° when viewed from 1 meter away (Cline, 1967), or 2.80° (or approximately 1mm) when viewed from 2 meters away (Gibson & Pick, 1963). But of all the possible directions others could be looking, one seems especially salient: When they are looking straight at us.

1.2.1 Eye Contact

Directed eyes grab us from the day we are born: Infants as young as 48 hours prefer to look at faces with a direct gaze (vs. an averted gaze; Farroni et al., 2002), suggesting that sensitivity to eye contact may be innate (a module called the 'Eye Direction Detector'; Baron-Cohen, 1997). But gazing direction is not just encoded: It also impacts the very way we process faces. For example, direct gaze facilitates identity recognition (Mason, Hood, & Macrae, 2004; Vuilleumier et al., 2005), gender discrimination (Macrae et al., 2002), and emotion perception (Adams & Kleck, 2003; cf. Bindemann, Burton, & Langton, 2008). Faces with direct gaze also appear to be more attractive (Kampe et al., 2001), likeable (Jones et al., 2006; Mason et al., 2005) , competent (Wheeler et al., 1979), dominant (Dovidio & Ellyson, 1982), among many other positive traits (for a review, see Kleinke, 1986). The pervasiveness of eye contact is apparent even just from the diverse domains of psychology that have explored its effects, from analyses of conversations (e.g., Kendon, 1967; Wang et al., 2011; Wohltjen & Wheatley, 2020), to studies of prosocial behavior (e.g., Bateson, Nettle, & Roberts, 2006; Nettle et al., 2013). And these effects have far reaching implications for our decisions, such as which candidates we choose to hire for a job (Forbes & Jackson, 1980), or which witnesses we trust in the courtroom (Hemsley & Doob, 1978; Neal & Brodsky, 2008). Indeed, law students are explicitly taught to encourage their witnesses to make eye contact with the jury during direct examination, and to block it during cross-examination (Brodsky, 2013).

But perhaps the most fundamental way in which eye contact affects us is simply that it grabs our attention: For example, faces staring at the observer (surrounded by faces looking away) are detected faster in visual search tasks than faces looking away (surrounded by faces staring at the observer)—a phenomenon known as the 'stare-in-the-crowd' effect (Senju, Hasegawa, & Tojo, 2005; von Grünau & Anston, 1995). And direct gaze is not just attention-grabbing, but downright distracting: The presence of irrelevant faces with direct (vs. averted) gaze impairs performance on a variety of tasks, from visual target detection (Senju & Hasegawa, 2005) and working memory (Nie et al., 2018; Wang & Apperly, 2017) to cognitive control (Conty et al., 2010), higher-level reasoning (Glenberg, Schroeder, & Robertson, 1998), language processing (Kajimura & Nomura, 2016), and spatial cognition (Buchanan et al., 2014; Markson & Paterson, 2009). And eye contact impacts not just our minds, but also our bodies, resulting in increased skin conductance (Hietanen et al., 2008; Pönkänen, Peltola, & Hietanen, 2011), heart rate (Akechi et al., 2013; Myllyneva & Hietanen, 2015), and facial muscle movements (Hietanen et al., 2018; for a review, see Hietanen, 2018).

In fact, direct gaze is so powerful that it impacts us even when it is not consciously perceived. This has been especially apparent in studies using continuous flash suppression (CFS)—a form of binocular rivalry in which a rapid stream of flashing masks shown to one eye temporarily renders invisible stimuli presented to the other (Tsuchiya & Koch, 2005; for a review, see Stein, 2019). In this work, faces looking at the observer break into awareness (escaping the interocular suppression from the flashing masks) faster than faces looking away (Chen & Yeh, 2012; Stein et al., 2011). This prioritization seems especially robust, since it has been demonstrated across various measures—e.g., not just breakthrough times, but also eye movements patterns (Rothkirch et al., 2015), and across diverse paradigms—e.g., not just in the absence of visual awareness, but also when the subject is aware, but their visual attention is occupied by a concomitant task (Yokoyama et al., 2014).

1.2.2 Gaze Cueing

If directed gaze is so evidently powerful in that is signals whether someone is looking towards us, perceiving others' gazing direction when their gaze is averted seems equally important, insofar as it signals where in the environment others are looking. The perceived direction of others' gaze is so salient that we tend to look in the same direction they are looking (a phenomenon known as 'gaze following'; Kuhn & Benson, 2007; Mansfield, Farroni, & Johnson, 2003; Ricciardelli et al., 2003). In fact, the perception of averted gaze produces not just overt eye and head movements as in gaze following, but also covert shifts of attention, as in when our attention automatically shifts in the direction that others are looking (a phenomenon known as 'gaze cueing'; e.g., Driver et al., 1999; for a review, see Frischen, Bayliss, & Tipper, 2007). This phenomenon is typically studied using variations of the spatial cueing paradigm (Posner, 1980), wherein observers are asked to fixate in the center of the screen and respond to a peripheral target; critically, their responses are affected by a face cue that appears at the center of the display prior to the target: They are faster when the face is looking in the same direction as the target ('congruent' cue), and slower when the face is looking in the opposite direction as the target ('incongruent' cue), and this difference is thought to reflect reflexive orienting in the direction that the face cue is looking.

Just like the perception of eye contact, gaze cueing of attention seems reflexive: It occurs rapidly (within about 100 ms of seeing a gaze cue), and disappears at longer intervals that allow observers to redistribute attention (Downing, Dodds, & Bray, 2004; Driver et al., 1999;

Friesen, Ristic, & Kingstone, 2004). It also seems to occur automatically, since it is not affected by secondary tasks (Hayward & Ristic, 2013; Law, Langton, & Logie, 2010; cf. Bobak & Langton, 2015), and involuntarily, since it persists not just when the gaze is nonpredictive of target location (e.g., Friesen & Kingstone, 1998), but also when counterpredictive (e.g., Driver et al., 1999). It also persists when observers are specifically instructed to ignore the face cue (e.g., Driver et al., 1999), or when they are informed about the future target location in advance (Galfano et al., 2012). While gaze cueing can be in part modulated by facial features such as dominance (e.g., Jones et al., 2010), trustworthiness (e.g., Petrican et al., 2013), or familiarity (e.g., Deaner, Shepherd, & Platt, 2007; for reviews, see Barbato, Almulla, & Marotta, 2020; Dalmaso, Castelli, & Galfano, 2020), in general it seems to occur with all sorts of stimuli resembling eyes: realistic faces (e.g., Hietanen & Leppänen, 2003; Langton & Bruce, 2000), virtual agents (e.g., Nuku & Bekkering, 2008), schematic faces (e.g., Friesen & Kingstone, 1998), animal faces (Quadflieg, Mason, & Macrae, 2004), objects that resemble faces (as in the so-called 'pareidolia' faces: Takahashi & Watanabe, 2013), and objects that do not even resemble faces but have eyes pasted on them (e.g., a glove; Quadflieg et al., 2004). And just like eye contact, gaze cueing can impact other cognitive processes: Objects that have been looked at by others are rated more favorably (Bayliss et al., 2006; King, Rowe, & Leonards, 2011), and faces that cue attention reliably are later rated as more trustworthy—despite no differences in the cueing effect itself (Bayliss & Tipper, 2006; Mattavelli et al., in press).

1.3 Why so Special?

The wealth of evidence reviewed so far converges on a clear conclusion: The eyes seem special. They are the main and richest feature we focus on when viewing others, and have great effects on us—from the powerful salience of eye contact to the irresistible effects of gaze cueing. One of the most striking aspects of these effects is that they suggest there might be specialized mechanisms built into our minds tuned to detect and process eyes, resulting in automatic mechanisms of perceptual analysis, attentional prioritization, and mnemonic biases. And the eyes also seem to be truly special in this respect compared to other physical features: There does not seem to be any nose-specific mechanism involved in processing others' noses, or any privileged mechanism involved in processing others' foreheads, while phenomena like the 'stare in the crowd' or gaze cueing of attention seem to reflect eye-specific mechanisms. So I ask: why might the eyes be so special?

1.3.1 Salient Eyes

The special status of others' eyes is typically taken to reflect their unique properties as visual stimuli. And this seems obvious, given that they are indeed extremely visually salient stimuli. First of all, they are positioned in an especially central and salient region of the face, framed by high cheekbones and sharp highbrows, and highlighted by the reduction in facial protrusion that has developed in hominoids but not in other apes (Emery, 2000). The reader also will not be surprised to learn that the human eye region comprises not one but two eyes; and because these are placed at a relatively short intra-ocular distance, another person's eyes are likely processed by both of the perceiver's eyes, resulting in redundancy after binocular fusion. Extant work has even suggested that infants' preferences for looking at the eyes can be reproduced in a model that simply posits that visual input from the two eyes that overlaps when binocularly fused is preferentially attended (Wilkinson et al., 2014).

But beyond their position within the face, the morphology of the eyes themselves makes them especially salient. Comparative studies assessing eye morphology in half of all extant primate species (Kobayashi & Kohshima, 1997, 2001; see Fig. 1.1 for examples) revealed that compared to other primates, the human eye is: (1) The only one with a white sclera, which heightens contrast with the dark pupil; (2) The only one with a sclera that is paler than both the skin and the iris, which facilitates eye detection and gaze direction discrimination; (3) The one with the largest exposed sclera size in the eye outline—up to three times more than other primates (Kaplan & Rogers, 2002), which increases the range of possible eyeball movements; and (4) The one with the largest width-to-height ratio of the eye outline, which allows for a larger range of horizontal displacement and facilitates gaze direction discrimination.

These morphological properties are hypothesized to have specifically evolved to support accurate gaze perception at the expense of camouflaging (the 'gaze enhancement hypothesis'; Kobayashi & Kohshima, 2001), and empirical evidence has shown that indeed perception of gazing direction relies on the amount of exposed sclera and its contrast with the pupil (Anstis, Mayhew, & Morley, 1969; cf. Symons et al., 2004). For example, increasing the luminance of the sclerae enhances gaze perception (Yorzinski & Miller, 2020), both during daytime and nighttime (Yorzinski, Harbourne, & Thompson, 2021). In the 'bloodshot' illusion (demonstrated in Fig. 1.2a), percepts of gazing direction can be altered simply by darkening the sclera on either side of the pupil (Ando, 2002), or the skin on either side of the eye (Ando, 2004). In fact, even just moving a lamp horizontally within a room changes perceived gaze direction (West, 2013). And in the 'Bogart' illusion (demonstrated in Fig. 1.2b), percepts of gazing direction can be reversed by inverting the polarity of the eyes (Ricciardelli, Baylis, & Driver, 2000; Sinha, 2000). The impact of luminance and contrast on perceived gaze direction has suggested that "the visual system follows an inflexible contrast rule for gaze perception, invariably treating the dark part of the eye image as the part that does the looking" (Ricciardelli et al., 2000, p. 12).

These powerful morphological features have been hypothesized to drive the effects others' gaze has on us. For example, the Eye Contact Detector module is thought to "track

Figure 1.1: External morphology of the eyes in primate species. The human eye seems exceptional compared to that of non-human primates in the bright coloration of the sclera, the sharp contrast with both the skin and the iris, the large extent of exposed sclera, and the extraordinary horizontal elongation of the eye outline.

Figure 1.2: The effects of eye coloration on perceived gaze direction. (a) The 'bloodshot' illusion: The same picture of a person facing forward appears to be gazing to her left and right, when her sclerae have been darkened on the left and right side of the iris, respectively (Ando, 2002). (b) The 'Bogart' illusion: The same picture of a person looking to her left appears to be looking right when the image colors are inverted, such that the sclera is now dark and the pupil is now light (Ricciardelli et al., 2000; Sinha, 2000).

and code the spatial position of the d[arker] region relative to the w[hite] region" (Baron-Cohen, 1994, p. 212), such that the perception of eye contact is triggered by a simple 'dark in center' rule. In other words, gaze is powerful because it is computed by a specialized "face detection system [...] using contrast information between the circular dark iris and the white sclera" (Palanica & Itier, 2011, p. 13). These salient visual features are also thought to underlie the phenomenon of gaze cueing, which is "driven by the mechanics of eye deviation perception" (Cole, Smith, & Atkinson, 2015, p. 1105), or "driven by a dedicated gaze perception mechanism that is sensitive to a unique property of the human eye (i.e. a darker iris surrounded by a wide white sclera)" (Ricciardelli et al., 2009, p. 355), such that "reflexive attentional orienting following an observed gaze is caused by the detection of the position of the iris, which is a darker region relative to the sclera" (Yoshizaki & Kato, 2011, p. 339). In sum, the effects of eye gaze seem to stem from eye-specific mechanisms, such that "the mere possession of eyelike stimuli is sufficient to trigger attentional orienting" (Quadflieg et al., 2004, p. 830).

1.4 A New Theory: Perceiving Perception

While vision science has typically taken the power of gaze to stem from the salience of the eyes per se, we suggest a new perspective: Rather than reflecting lower-level visual properties, the effects of gaze reflect the apprehension of higher-level properties of the *minds* behind the eyes—namely the intentions and attention of other people. Coincidentally, these states can sometimes be signaled by the eyes: When choosing among multiple options, people tend to look towards the objects they desire most (Shimojo et al., 2003; Simion & Shimojo, 2006); when making tea or a sandwich, people look towards the ingredients they will grab next (Land et al., 1999; Land & Hayhoe, 2001); when walking along a trail, people look towards the locations they will step on next (Hollands, Patla, & Vickers, 2002; Matthis, Yates, & Hayhoe, 2018). As such, the eyes are not just visual cues—they are some of the most direct cues we have to others' minds, and we argue that it is the perception of these deeper properties that drives their powerful effects.

Imagine you are at a crosswalk: You look both ways before crossing, and see a car approaching. Suddenly you notice the driver is looking right at you. Did they notice you? Should you cross? You can certainly reason about this and come up with an answer: If of all possible locations they are looking exactly in your direction, it is likely you are salient in their mind; drivers are also legally and morally required to look out for pedestrians at crosswalks, so it is likely they noticed you. By this time, they are honking: Why haven't you crossed yet? I argue that the impression that someone is seeing us is not something we need to reason about: It is something we extract rapidly, effortlessly, and automatically. In short, others' attention can be *seen*.

This intuition is based on recent work suggesting that perception might recover not just the physical structure of our environment (e.g., colors, shapes, or motion), but also its social structure (for reviews, see Scholl & Gao, 2013; Scholl & Tremoulet, 2000). For example, imagine a display containing numerous darts and one disc; the darts are moving around in a random fashion, but they are always pointing towards the disc. Upon seeing such displays, observers typically report a distinct impression that the darts are actively pursuing the circle (the 'wolfpack effect'; Gao, McCarthy, & Scholl, 2010). These attributions of animacy and intentionality to simple shapes are powerful: They hold across cultures (Barrett et al., 2005), and arise early in development (Csibra, 2008; Rochat, Morgan, & Carpenter, 1997), but most importantly they exhibit classic signatures of perceptual processing. For example, they result in a distinct phenomenology, despite observers of course knowing a dart cannot have a mental state (Gao et al., 2010); they depend on subtle features of the visual displays (e.g., the orientation of the darts; Gao, Newman, & Scholl, 2009); they influence us even when irrelevant to the task at hand (van Buren, Uddenberg, & Scholl, 2016); and they interact with other visual properties such as objecthood (van Buren, Gao, & Scholl, 2017; for a review, see Scholl & Gao, 2013).

This suggests that others' attention is not just inferred or represented: It is automatically extracted in perception, becoming integrated in our own visual experience of the world. If the direction of others' attention can be extracted in perception, then it seems that the reflexive nature of gaze perception need not necessarily reflect an effect of the eyes as visual stimuli, but rather an effect of the eyes as a proximal cue to others' attention. In short, effects of 'eye contact' might be better characterized as but one case of more general effects of 'mind contact'.

1.5 Overview of the Current Chapters

This framework makes two complementary predictions, which are empirically tested in Chapters 2-5. First, if gaze is a reliable indicator of where our attention is directed, it is certainly not the only one: Our attention can be indicated in numerous other ways, e.g., body orientation, hand gestures, etc. In fact, even inanimate objects can appear to be intentional on the basis of spatial or motion cues (Heider & Simmel, 1944; for a review, see Scholl & Gao, 2013). If the effects of eye gaze are truly driven by the perception of minds, they should arise from such stimuli that signify others' attentional states, even if they have no resemblance to eyes. This prediction is empirically tested in studies of 'minds without eyes' (Chapter 2 and Chapter 3), where notorious effects of eye gaze (such as the stare-in-the-crowd effect) replicate with eye-less stimuli that are nonetheless perceived as being directed toward or away from the observer. In particular, we demonstrate that simple shapes can induce putatively eye-specific effects as in the 'distracting' effect of eye contact (Chapter 2), and the attentional prioritization of 'stares in the crowd' (Chapter 3).

Second, if the eyes are a reliable indicator of where our attention is directed, this is certainly not always the case: Sometimes we attend covertly (i.e., shift our attentional focus despite maintaining fixation; Cameron, Tai, & Carrasco, 2002; Posner, 1980). And conversely, we might fail to attend to information despite looking right at it (Mack, 2003; Mack & Rock, 1998); we might attend to some but not other objects within the same area (Scholl, 2001), and to some but not other features within the same object (e.g., Saenz, Buracas, & Boynton, 2002). If the effects of eye gaze are truly driven by the perception of minds, they should be reduced when the eyes do not signify any underlying pattern of attention or intention. This prediction is empirically tested in studies of 'eyes without minds' (Chapter 4 and Chapter 5), where powerful effects of the eyes (such as gaze cueing) are reduced when they do not actually signify underlying patterns of attention and intention, even if they clearly look like eyes. In particular, we demonstrate that gaze cueing is reduced in the phenomenon we have dubbed ‘gaze deflection’ (Chapter 4), and that unconscious prioritization of attention is reduced when the faces seem inattentive, in a novel form of ‘unconscious pupillometry’ (Chapter 5).

These chapters were written as individual journal articles, and they are all published (Colombatto, Chen, & Scholl, 2020; Colombatto, van Buren, & Scholl, 2019, 2020) or in press (Colombatto & Scholl, in press). As a result, they stand or their own and can be read independently, in any order.

2

Intentionally distracting:

Working memory is disrupted
by the perception of other agents attending to you
—even without eye-gaze cues

2.1 Abstract

Of all the visual stimuli you can perceive, perhaps the most important are other people's eyes. And this is especially true when those eyes are looking at you: direct gaze has profound influences, even at the level of basic cognitive processes such as working memory. For example, memory for the properties of simple geometric shapes is disrupted by the presence of other eyes gazing at you. But are such effects really specific to direct gaze per se? See-

ing eyes is undoubtedly important, but presumably only because of what it tells us about the "mind behind the eyes"—i.e., about others' attention and intentions. This suggests that the same effects might arise even without eyes, as long as an agent's directed attention is conveyed by other means. Here we tested the impact on working memory of simple "mouth" shapes—which in no way resemble eyes, yet can still be readily seen as intentionally facing you (or not). Just as with gaze cues, the ability to detect changes in geometric shapes was impaired by direct (compared to averted) mouths—but not in very similar control stimuli that were not perceived as intentional. We conclude that this disruption of working memory reflects a general phenomenon of "mind contact," rather than a specific effect of eye contact.

2.2 Introduction

The most salient stimuli we encounter in everyday life are arguably eyes: we constantly monitor where others are looking (for reviews, see Emery, 2000; Grossmann, 2017; Langton et al., 2000), and when exploring others' faces we attend most to the eye region (e.g., Henderson et al., 2005; Janik et al., 1978). This is understandable, given that the eyes are exceptionally reliable cues for deciphering identity (Peterson & Eckstein, 2012; Schyns et al., 2002), demographics (Macrae et al., 2002; Provine et al., 2013; Russell et al., 2014), emotions (Ekman & Friesen, 1971; for a review, see Itier & Batty, 2009), and even character traits such as competence (Wheeler et al., 1979) and dominance (Dovidio & Ellyson, 1982).

But perhaps the most obvious way in which eyes are informative is that they indicate where in the environment people are looking, and they signal others' intentions—and most importantly, they can indicate when others are attending to (and perhaps have intentions that concern) *us*. In fact, eye contact is preferentially attended from the very beginning of life (e.g., Farroni et al., 2002), and it can draw attention even when it is not consciously perceived (Chen

& Yeh, 2012; Stein et al., 2011). Moreover, the cognitive processing of faces is greatly impacted by how the eyes are directed, in contexts ranging from long-term memory (e.g., Mason et al., 2004) to aesthetic experience (e.g., Chen et al., 2018). But perhaps the clearest example of the power of the eyes is how they can also influence the processing of *other* (eye-less) objects in a scene.

2.2.1 Direct Gaze, Distraction, and Working Memory

One of the most robust effects of the eyes is that direct gaze is distracting. For example, when discriminating the colors of words in a Stroop task, performance is impaired if the words are accompanied by faces looking at us (vs. faces with closed eyes; Conty et al., 2010). This distracting power of eye contact has been demonstrated in a variety of contexts, including simple visual target detection (Senju & Hasegawa, 2005), higher-level reasoning (Glenberg et al., 1998), language processing (Kajimura & Nomura, 2016), and spatial cognition (Buchananet al., 2014; Markson & Paterson, 2009). And conversely, looking away from others (e.g., staring at the ceiling) facilitates knowledge retrieval and concept learning in adults (Glenberg et al., 1998) and children (Doherty-Sneddon et al., 2001; Phelps et al., 2006), and even in atypical development (Riby et al., 2012).

This influence of direct gaze is especially apparent when considering how eye contact influences working memory for other objects in a scene. When asked to detect changes between two consecutive arrays of geometric shapes (for example when one shape changes from a circle to a hexagon), performance is impaired by the presence of (utterly task irrelevant) eyes looking at us (vs. looking away, or at one of the other shapes; Nie et al., 2018; Wang & Apperly, 2017).

In general, these far-reaching influences of direct gaze on seeing and thinking have been taken as a testament to “the special status of eye contact and mutual gaze in social situations” (Buchanan et al., 2014, p. 5), revealing its power, but also its uniqueness. For example, working memory disruptions have been interpreted to suggest that “the mere presence of direct gaze automatically calls for processing resources [...], at the expense of any concurrent visual processing outside the facial area” (Conty et al., 2010, p. 134), and that “although many directional cues might trigger reflexive shifts of attention [...], gaze cues are more strongly [influential to] internal object representations [...], possibly because they access a neural architecture that is specialized for processing gaze direction” (Nie et al., 2018, p. 93).

2.2.2 The Current Studies: Distracting Eyes, or Distracting Minds?

While this previous work clearly demonstrates the power of perceived eye gaze, here we ask whether these effects must really be eye-specific. Might they instead reflect responses to a deeper property that the eyes (but not only the eyes) reliably signal—namely the direction of other agents’ attention and intentions? Eye gaze predicts which action someone is going to perform next in a sequence of tasks (Land & Hayhoe, 2001), where their attention is located in conversations (Foulsham et al., 2010), and which objects they desire (King et al., 2011), etc. In this way, perhaps the eyes are important because they are informative about others’ minds. If these effects reflect the “special status of eye contact” as a visual stimulus (Buchanan et al., 2014, p. 5), as is commonly assumed, then they should obviously require the presence of eye-like stimuli in the first place. But if these effects instead reflect the perception of others’ *minds* (e.g., their underlying patterns of attention and intentions), then they should also be triggered by stimuli that don’t resemble eyes at all, as long as the agents’ attention is signaled by other means.

Here we directly tested these competing predictions by asking whether the very same distracting effects would arise for simple "mouth" stimuli that look nothing like eyes, yet are readily seen as facing towards or away from the observer—as depicted in Fig. 2.1. In particular, we followed the procedure of Wang and Apperly (2017) exactly, but substituted direct and averted mouth stimuli (as in Fig. 2.2b) for their direct and averted gaze stimuli (as in Fig. 2.2a). Would this alternate means of conveying directed attention still impair visual working memory for the other properties of objects in the scene?

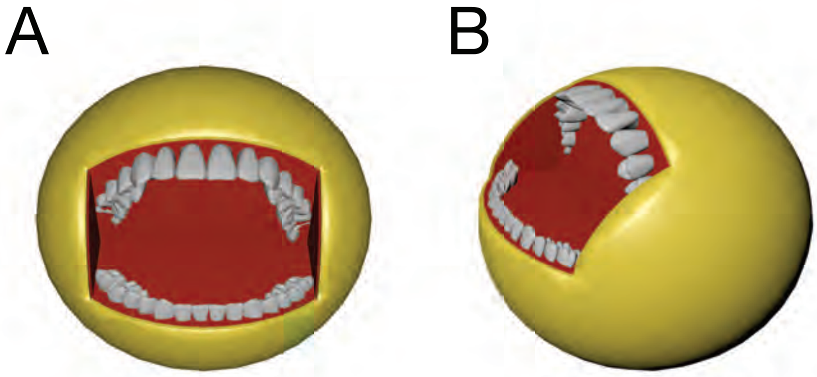

Figure 2.1: Examples of the mouth stimuli used in Experiments 1 and 2.

2.3 EXPERIMENT 1: DISTRACTING MOUTHS AND MINDS

Following Wang and Apperly (2017, Experiment 1a), observers viewed briefly presented pairs of displays (one after the other) containing direct or averted mouths, and simply had to detect whether one of the shapes had changed its color or shape between the two presentations.

2.3.1 METHOD

OBSERVERS

Sixteen members of the Yale community (13 females; average age = 21.00 years, *SD* = 3.01 years) participated in exchange for monetary compensation. (This sample size was chosen ahead of time to exactly match that of Wang & Apperly, 2017.)

APPARATUS

Stimuli were presented on a Dell 1905FP monitor with a 60-Hz refresh rate, using custom software written in Python with the PsychoPy libraries (Peirce, 2007). Observers sat in a dimly lit room without restraint approximately 60 cm away from the display, which subtended 34.87° × 28.21° (with all visual extents reported below computed based on this viewing distance).

STIMULI

The mouths were generated using Blender (version 2.76). Each mouth consisted of a realistic 3D model of human teeth embedded in a sphere and could face one of five directions: straight ahead (for mouths directed straight at the observer, as in Fig. 2.1b), or oriented 45°, 135°, 225°, or 315° within the image plane (for mouths directed away from the observer, as in Fig. 2.1a). The color of the sphere was varied to obtain six different mouths (yellow, orange, pink, purple, light blue, and green), with white teeth and a red inside.

Displays included either three or four mouths placed in random non-overlapping locations on a white background (each at least 1.32° from the nearest display border), and an equal number of gray geometric shapes (randomly chosen from a triangle, square, diamond, trapezoid, hexagon, and circle), each placed diagonally from a mouth (top-left, top-right, bottom-left, or bottom-right in an imaginary grid) at a distance randomly jittered between 1.41° and 2.81°. The colors of the mouths and the shapes of the gray geometric figures were randomly chosen such that no color or shape appeared more than once in any given display. The mouths in each trial either faced the observer (as in the top panel of Fig. 2.2b; "Directed-at-You") or faced their respective shapes (as in the middle panel of Fig. 2.2b; "Directed-at-Shapes")—with the same spatial arrangements, colors, and shapes used in each case.

To construct the displays with changes, each of the initial scenes was modified in two ways. In *Shape* changes, a randomly selected geometric shape was replaced with a different shape (presented in the same location) that was not already present in the display. In *Mouth* changes, a randomly selected mouth appeared in a different randomly selected color that was not already present in the display. The same change was always made to both a Directed-at-You display and its matched Directed-at-Shapes display.

A central black bounding frame (15.37° × 12.60°, drawn with a stroke of .06°) was present throughout each entire trial to mark the active region of the display, along with two letter strings that served as reminders for the response key mapping (presented below the bounding box, with the highest point of the tallest letter 8.08° below the center of the display): "Change" (presented on the left, with its left edge 7.69° from the display's center) and "No Change" (presented on the right, with its left edge 3.43° from the display's center).

PROCEDURE

Each trial began with a central black fixation cross (0.59° × 0.59°) for 1 s, followed by the first display (13.99° × 11.21°) for 100 ms. After a 900-ms blank interval, a second display was presented and remained visible until a response was made. (Within these displays, the mouths each subtended 2.70° × 2.45°, and the shapes each subtended 1.91° × 1.91°—except for the diamond [2.14° × 2.14°] and the hexagon [2.19° × 1.91°].) Observers were instructed to indicate whether a change had occurred by pressing one of the two arrow keys, and the next trial started after a 250-ms blank delay following each response.

Observers completed 400 trials: 25 random spatial arrangements × 2 directions of attention (Directed-at-You, Directed-at-Shapes) × 2 set sizes (3, 4) × 2 possible outcomes (Change, No Change) × 2 repetitions. These trials were presented in random order, split into four blocks of 100 trials each, presented in a random block order. Two of the blocks featured shape identity changes, and two featured mouth color changes. The first four trials of each block were treated as practice trials, data for which were not recorded.

2.3.2 Results and Discussion

We categorized each response as a hit, miss, false alarm, or correct rejection, and then computed *d'* (a measure of sensitivity, as distinct from response bias; Green & Swets, 1966) for all conditions. All observers were within 2 standard deviations of the mean sensitivity in all conditions, and hence all were included in the analyses (following Wang & Apperly, 2017). The *d'* scores for the Directed-at-You and Directed-at-Shapes conditions are depicted in Fig. 2.2b, and inspection of this figure reveals a reliable impairment in change detection performance for Directed-at-You versus Directed-at-Shapes displays (1.47 vs. 1.62, $t(15) = 2.73$, $p = .015$, $d = .28$)—a difference analogous to that observed by Wang and Apperly (2017) using direct versus averted eye gaze (as depicted in Fig. 2.2a)[1]. Thus, the impairment of visual working memory

---

[1] Following Wang and Apperly (2017), we also analyzed our data using a 2 (Mouth Direction) × 2 (Change Type) × 2 (Set Size) repeated-measures analysis of variance (ANOVA) with change detection sensitivity as the dependent variable. This yielded a main effect of mouth direction (with changes in Directed-at-You displays detected less accurately than changes in Directed-at-Shapes displays; $F(1, 15) = 4.69$, $p = .047$, $\eta_p^2 = .24$), a main

by direct gaze seems not to require *gaze*, per se, as long as directed attention and intentions are depicted in other ways.

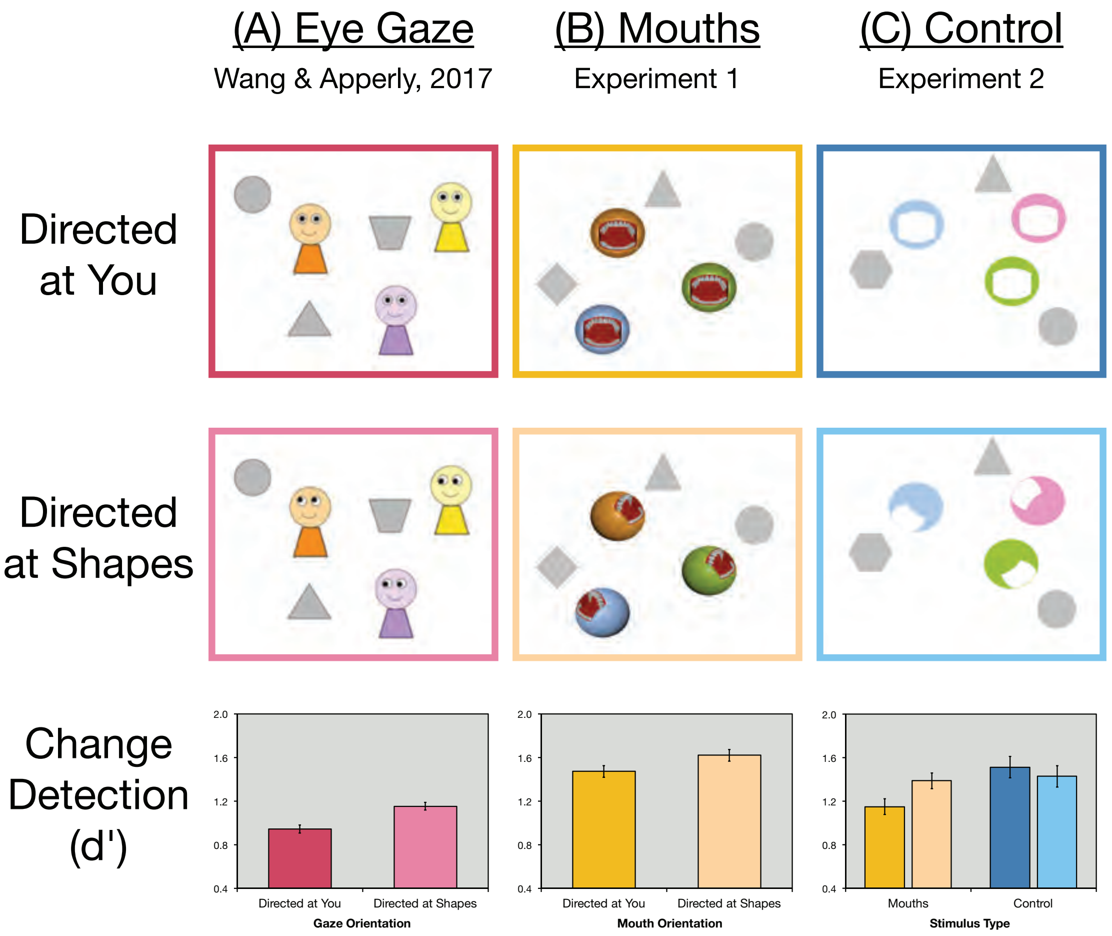

Figure 2.2: Working memory disruptions caused by eye gaze, mouths, and control stimuli. (a) Sample displays and results from Experiment 1a of Wang and Apperly (2017). (b) Sample displays and results from Experiment 1. (c) Sample displays and results from Experiment 2. Error bars reflect 95 % confidence intervals, subtracting out the shared variance.

---

effect of change type (with changes to shapes detected less accurately than changes to mouths; $F(1, 15) = 42.79$, $p < .001$, $\eta_p^2 = .74$), and a main effect of set size (with changes in displays with four mouths and four shapes detected less accurately than changes in displays with three mouths and three shapes; $F(1, 15) = 121.19$, $p < .001$, $\eta_p^2 = .89$), but no significant interactions (all $Fs < 2$).

2.4 EXPERIMENT 2: DIRECT REPLICATION + NON-AGENTIAL CONTROL STIMULI

We interpret the results of Experiment 1 in terms of a novel type of mouth-induced social attention: the mouths themselves viscerally indicated the presence of agents, along with those agents' directions of attention and intentions—despite the lack of eyes. However, beyond any appeal to perceived agency, our mouth stimuli also had a simple visual asymmetry, with a smaller part of the sphere (i.e., the open mouth) clearly presented either centrally or to one side. And correspondingly, the Directed-at-You spheres had less visible color than did the Directed-at-Shapes spheres.

To ensure that our results were due to the perceived agency of the mouths rather than these lower-level visual properties, we ran a direct replication of Experiment 1, along with an added between-subjects factor: for half of the observers, the entire mouth region simply shared the color of the background. As can be appreciated in Fig. 2.2c, this manipulation eliminated any percept of mouths or agents, while retaining the same differential symmetry and degree of visible color. We predicted that the results of Experiment 1 would replicate with mouths, but not with these non-agential control displays.

2.4.1 METHOD

This experiment was identical to Experiment 1, except as noted here. The sample size was doubled (to 32; 23 females; average age = 22.09 years, *SD* = 3.67 years) to maintain the same number of observers per cell as in both Experiment 1 and Wang and Apperly (2017). Half of the observers completed a direct replication of Experiment 1, and the other half completed a replication with control stimuli. Control stimuli were generated using the same criteria as the mouths, except that (1) they were rendered with a more luminous light source, such that the color of the sphere would be uniform (thus effectively removing any depth information), and (2) the cutout of the sphere (where the teeth were placed in the mouths) was drawn in solid white (thus effectively removing any trace of the mouth).

2.4.2 Results and Discussion

The average change detection sensitivities for Directed-at-You and Directed-at-Shapes displays are depicted separately for the mouth and control stimuli in Fig. 2.2c. Inspection of this figure suggests two clear patterns. First, the mouth condition replicated the impairment for Directed-at-You displays that was observed in Experiment 1. Second, no such effect occurred for the control stimuli (which, if anything, trended in the opposite direction). These impressions were verified with a 2 (stimulus type: mouths vs. control) × 2 (direction: Directed-at-You vs. Directed-at-Shapes) mixed analysis of variance, which revealed no effect of stimulus type ($F(1, 30) = 0.75$, $p = .393$, $\eta_p^2 = .02$), no effect of stimulus direction ($F(1, 30) = 1.55$, $p = .222$, $\eta_p^2 = .05$), and—most importantly—a highly reliable interaction between these factors ($F(1, 30) = 6.68$, $p = .015$, $\eta_p^2 = .18$). Specific comparisons then confirmed that observers in the mouth condition were again less sensitive to changes in Directed-at-You displays compared to changes in Directed-at-Shapes displays (1.15 vs. 1.39, $t(15) = 3.25$, $p = .005$, $d = .42$), but that no such difference occurred with the control shapes (1.51 vs. 1.43, $t(15) = 0.83$, $p = .420$, $d = .11$).

Beyond demonstrating the strength and replicability of the primary effect, these results indicate that the impairment of visual working memory by Directed-at-You mouths is due to the perception of the mouths as directed agents.

2.5 General Discussion

The primary result of this study (replicated in both Experiment 1 and Experiment 2) was extremely clear: visual working memory for the details of displays is impaired not only by the presence of eyes that are directly looking at you (as in Wang & Apperly, 2017), but also by the presence of *mouths* that are directly facing you. (By design, these mouth stimuli themselves bore no resemblance to eyes—though of course they may have led observers to effectively "fill in" other features, such as eyes, that are associated with agents.) Critically, this effect seems to depend on the perceived agency of the mouths, since it vanished with nearly identical control stimuli that are not perceived as intentional. We conclude that the disruption of visual working memory by direct gaze is not specific to gaze after all: these results reflect not a specific phenomenon of eye contact, but rather a more general phenomenon of "mind contact."

Of course, the present study did not attempt to directly compare the magnitudes of the effects with eyes versus mouths. It is difficult to make predictions about such comparisons based on previous work, in part because the eyes are typically contrasted with every other part of the face at once, rather than with other particular features (e.g., Gilad et al., 2009; Itier et al., 2006). And when the eyes have been directly contrasted with other features (e.g., noses), these comparisons have typically not been in the context of averted features that may signal the direction of attention (e.g., Looser & Wheatley, 2010). (Of course, different facial features may be more or less important for communicating other information such as emotion; e.g., Eisenbarth & Alpers, 2011. But that needn't have any consequences for whether the eyes are special in terms of directing or distracting attention and memory.) We suspect, based on the "mind contact" framework, that comparisons between features such as eyes and mouths might depend not on these stimulus categories themselves, but rather on how effectively a given stimulus conveys an agent's attention or intentions. As a result, many eye stimuli may be more effective than many mouth stimuli, but the reverse could also be true in some circumstances. This perspective also suggests that similar effects might be possible with some other sorts of eyeless stimuli such as pointing fingers—but perhaps not with other non-agential stimuli, such as arrows.

The current work thus integrates the vast literature on face perception with the still largely unconnected literature on the perception of animacy and intentionality. The key distinction in these experiments between superficial surface features (i.e., the eyes themselves) and the deeper properties they signify (i.e., the perceived direction of attention and intentions) was in fact inspired by research demonstrating that even simple (and eye-less) geometric shapes are readily seen as alive and goal-directed when they move in certain ways (Heider & Simmel, 1944; Michotte, 1950/1991; for reviews, see Scholl & Gao, 2013; Scholl & Tremoulet, 2000). Just as in the case of eye contact, sensitivity to these simple cues to animacy arises early in development (e.g., Gergely et al., 1995; Southgate & Csibra, 2009) and has been documented in disparate cultures (Barrett et al., 2005). And interestingly, perceived animacy also influences a variety of downstream processes such as attention (Gao et al., 2018; Meyerhoff et al., 2013), spatial memory (van Buren & Scholl, 2017) and visuomotor behavior (Gao et al., 2010; van Buren et al., 2016). Our results thus add to a growing recognition that our minds are especially well tuned to extracting intentionality in our surroundings, and they offer a new perspective on eye contact as a special case of perceived intentionality that we call "mind contact."

3

Gazing without eyes:

A "stare-in-the-crowd" effect induced by simple geometric shapes

3.1 ABSTRACT

OF ALL THE VISUAL STIMULI YOU CAN PERCEIVE, Of the many effects that eye contact has, perhaps the most powerful is the *stare-in-the-crowd effect*, wherein faces are detected more readily when they look directly toward you. This is commonly attributed to others' eyes being especially salient visual stimuli, but here we ask whether stares-in-the-crowd might arise instead from a deeper property that the eyes (but not only the eyes) signify: the direction of others' attention and intentions. In fact, even simple geometric shapes can be seen as intentional, as when numerous randomly scattered cones are all consistently pointing at *you*.

Accordingly, we show here that cones directed at the observer are detected faster (in fields of averted cones) than are cones averted away from the observer (in fields of directed cones). These results suggest that perceived intentionality itself captures attention—and that even in the absence of eyes, others' directed attention stands out in a crowd.

3.2 Introduction

One of the most striking things that we can perceive is the direct gaze of another person—as when we are eating in a restaurant and suddenly notice that someone across the room is staring right at us (yikes!). Such direct gaze cues have profound influences on many aspects of our mental lives (for a review see Emery, 2000). For example, we remember people better when they are looking at us (e.g., Mason et al., 2004); we see faces making eye contact as more competent (e.g., Wheeler et al., 1979); and faces with direct gaze also elicit increased mimicry during conversations (e.g., Wang et al., 2011). But perhaps the most fundamental way in which eye contact influences us is simply that we detect it so readily in the first place: direct gaze (in contrast to averted gaze) captures our attention (e.g., Böckler et al., 2014; Miyazaki et al., 2012; Senju & Hasegawa, 2005) from the very first few days of life (e.g., Farroni et al., 2002), and even when it is not consciously perceived (e.g., Chen & Yeh, 2012; Rothkirch et al., 2015; Stein et al., 2011).

Perhaps the most powerful demonstration of attention capture by direct gaze is the *stare-in-the-crowd effect*, wherein a face staring at the observer (surrounded by faces looking away) is detected faster than a face looking away (surrounded by faces staring at the observer; Conty et al., 2006; Doi & Ueda, 2007; Doi et al., 2009; Framorando et al., 2016; Palanica & Itier, 2011; Shirama, 2012; von Grünau & Anston, 1995; cf. Cooper et al., 2013). This visual-search advantage for *stares* has been attributed to "special physiological mechanisms for the processing of straight-gaze stimuli" (von Grünau & Anston, 1995, p. 1312) or a "face detection system [...] using contrast information between the circular dark iris and the white sclera" (Palanica & Itier, 2011, p. 13).

Direct gaze is important, however, not because of what it signals about the eyes per se, but because of what it signals about the deeper properties of other agents: where they are attending, and perhaps what their intentions are. For example, we typically look at people we are listening to (Foulsham et al., 2010), or objects we desire (King et al., 2011), or the locations we are about to act on (Ballard et al., 1997). So might the stare-in-the-crowd effect not be specific to the eyes after all? Might it instead reflect a more general effect on the efficiency of visual search of a broader class of cues related to agents' attention and intentions?

Here we employed a standard stare-in-the-crowd task (Senju et al., 2005; Experiment 1), but we replaced the faces with directed *cone* stimuli that in no way resemble eyes—but which are nonetheless readily seen as facing toward or away from the observer (see Fig. 3.1a and b). Of course it might initially seem odd to talk about directed cones as reflecting intentions, since after all one can easily tell from their shapes that they are not biological entities at all. Nevertheless, we speculate that the coordinated orientations of many such stimuli may in fact serve as a powerful cue to the presence of agency. Imagine walking down a forest trail and suddenly noticing that a single tree branch is pointing right at you; this would in no way signal the presence of an agent, since it could just be a coincidence. (In a dense forest with many branches, surely *some* just happen to be pointing at you.) Similarly, imagine that you come upon some flowers all of whose stems are pointing in the same direction (e.g., toward the east); this also would not suggest the presence of an agent, since these uniform orientations might just reflect the operation of an external factor (perhaps a strong wind). But now imagine walking in the forest and suddenly noticing that while all of the branches on the surrounding trees are pointing in different absolute directions, they are nevertheless all pointing directly at *you*. This sort of stimulus cannot be due to a simple coincidence; such arrangements thus signal the presence or action of an agent, and have several downstream effects that are *social* in nature (e.g., Gao et al., 2010; Takahashi et al., 2013). (This principle is especially apparent in the sculptures of the renowned British land artist Andy Goldsworthy, who re-arranges materials found in natural environments in a way that readily suggests the presence of a designer; see for examples "Hole covered with small pointed rocks" `https://www.goldsworthy.cc.gla.ac.uk/image/?id=ag_01380&t=1` or "Woven branch circular arch" `https://www.goldsworthy.cc.gla.ac.uk/image/?id=ag_03744&t=1`. Indeed, one of the factors that makes such artwork so striking is the stark contrast between the nonagentic materials and the agentic arrangements.)

Given that coordinated-orientation stimuli may signal the presence of agency, we predicted that such stimuli might also give rise to stare-in-the-crowd effects—despite the absence of eyes.

3.3 Experiment 1a: Staring Without Eyes?

Following Senju et al. (2005, Experiment 1), observers viewed circular arrays of cones, and simply had to detect either an averted cone (in a field of directed cones) or a directed cone (in a field of averted cones), as depicted in Fig. 3.1a and b, respectively.

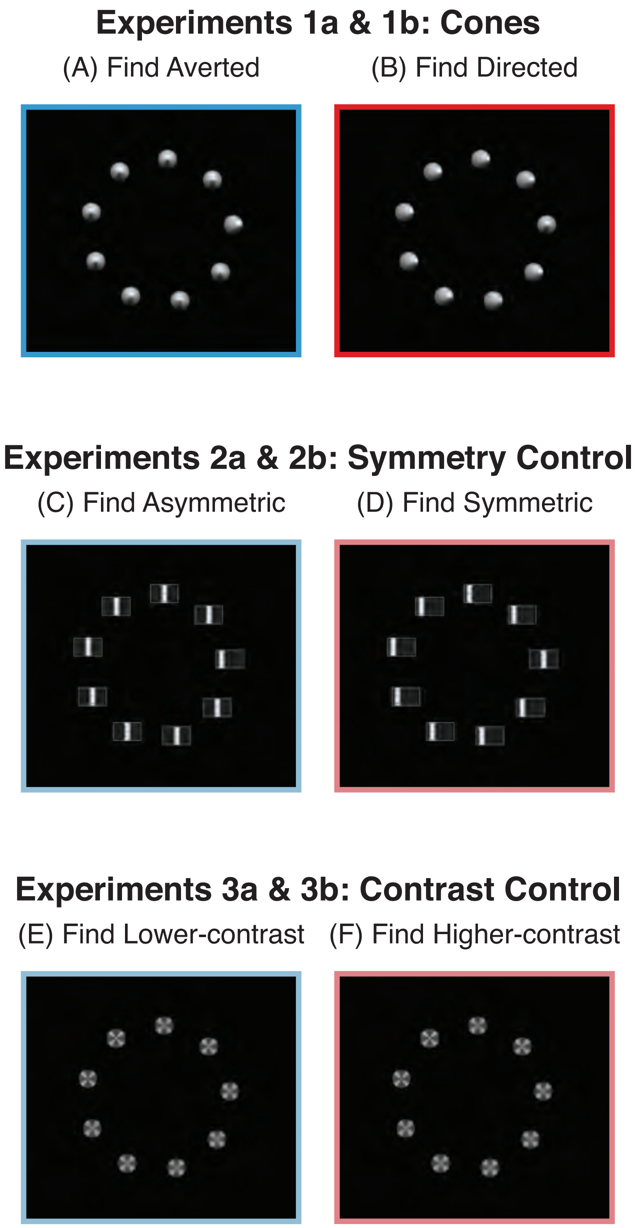

Figure 3.1: Sample displays from Target-Present trials of each experiment. (a) A sample display from Experiments 1a and 1b in which observers must find an averted target in a field of directed cones. (b) A sample display from Experiments 1a and 1b in which observers must find a directed target in a field of averted cones. (c) A sample display from Experiments 2a and 2b in which observers must find an asymmetric target in a field of symmetric poles. (d) A sample display from Experiments 2a and 2b in which observers must find a symmetric target in a field of asymmetric poles. (e) A sample display from Experiments 3a and 3b in which observers must find a lower contrast target in a field of higher contrast pinwheels. (f) A sample display from Experiments 3a and 3b in which observers must find a higher contrast target in a field of lower contrast pinwheels.

3.3.1 METHOD

OBSERVERS

Ten members of the Yale community (with an average age of 22.4 years) participated in exchange for monetary compensation. This sample size was determined before data collection began (arbitrarily rounded up from the sample of eight in Senju et al., 2005) and was fixed to be identical for each of the four studies reported here.

APPARATUS

Stimuli were presented on a Dell M992 CRT monitor with a 75 Hz refresh rate, using custom software written in Python with the PsychoPy libraries (Peirce, 2007). Observers sat in a dimly lit room without restraint approximately 60 cm from the display, which subtended 33.57° × 25.49°; all visual extents reported below were computed based on this viewing distance.

STIMULI AND PROCEDURE

Each trial began with a central white fixation cross (0.60° × 0.60°) on a black background for 500 ms, followed by a display depicting an array of gray cones (see Fig. 3.1a and b). Each cone was rendered in Blender (Blender Foundation, version 2.76) with simulated point lighting from above and a simulated camera directly in front of it, with the resulting stimuli matched for mean luminance using the SHINE toolbox in MATLAB (MathWorks, version R2017a). As depicted in Fig. 3.1a, directed cones thus had their points directly in the middle of their bounding circular bases, with a three-dimensional appearance due to the simulated lighting.

These cones were then rotated in Blender such that the resulting Averted stimuli (see Fig. 3.1b) had their points directly on the perimeter of their bounding circular bases. Each display included either five or nine cones subtending 1.94° × 1.94° each, and arranged on equidistant points of an imaginary circle (centered in the display) with a diameter of 16.73° (with one of the cones always present at 45°).

On half of trials, observers searched for an averted cone among directed cones (Find Averted; Fig. 3.1a), and on the other half of trials they searched for a directed cone among averted cones (Find Directed; Fig. 3.1b)—with the target present on half of the trials of each type, and the target absent on the remaining half (i.e., with all cones directed in Find Averted trials, and all cones averted in Find Directed trials). Observers indicated whether the target was present or not by pressing one of two keys. Upon response, feedback then appeared on the screen ("Good job!" or "—") for 500 ms—and if a response was not made within 1.5s, the trial ended, and a display appeared (for 5s) reminding them to respond faster. The next trial then started after a 1 s blank delay.

Following Senju et al. (2005), each observer completed four blocks of trials (2 Target Orientations [Find Averted, Find Directed] × 2 Averted Directions [left, right]), presented in a different random order for each observer, and with the relevant target displayed prior to the beginning of each block. Each block consisted of 40 trials (2 Array Sizes [5/9] × 2 Correct Responses [present/absent] × 10 Repetitions), presented in a different random order for each observer. Observers completed four blocks (one of each type) of 20 practice trials each prior to the start of the experiment (in order to acquaint them with the task, the key mappings, the time pressure, and all trial types), and they completed eight additional practice trials at the beginning of each block (to entrain them to the new target)—the data for which were not recorded.

3.3.2 Results and Discussion

Trials in which observers failed to respond, responded inaccurately, or responded in less than 100 ms were discarded (with these exclusion criteria adopted directly from Senju et al., 2005). (This resulted in an average number of analyzed trials of 151.40 out of the possible 160.) As depicted in Fig. 3.2a, responses were faster for Find Directed displays compared to Find Averted displays (0.64 vs. 0.75 s, $t(9) = 3.85$, $p = .004$, $d = 0.79$). We also categorized each response (excluding trials in which observers failed to respond) as a hit, miss, false alarm, or correct rejection, and then computed *d'* (a measure of sensitivity, as distinct from response bias; Green & Swets, 1966) for all conditions. As depicted in Fig. 3.2c, sensitivity was higher for Find Directed displays compared to Find Averted displays (3.76 vs. 3.27, $t(9) = 2.58$, $p = .030$, $d = 0.74$). These initial results suggest that the stare-in-the-crowd effect may indeed not be specific to eyes.

3.4 Experiment 1b: Direct Replication

Given the importance of direct replications, we reran the experiment on another group of 10 observers ($M_{age} = 22.1$), analyzing an average of 150.70 nonexcluded trials per observer. As depicted in Fig. 3.2b, responses were again faster for Find Directed (compared to Find Averted) displays (0.71 vs. 0.79 s, $t(9) = 2.46$, $p = .036$, $d = 0.68$). And as depicted in Fig. 3.2d, sensitivity was again higher for Find Directed (compared to Find Averted) displays (4.01 vs. 3.28, $t(9) = 3.46$, $p = .007$, $d = 1.05$).

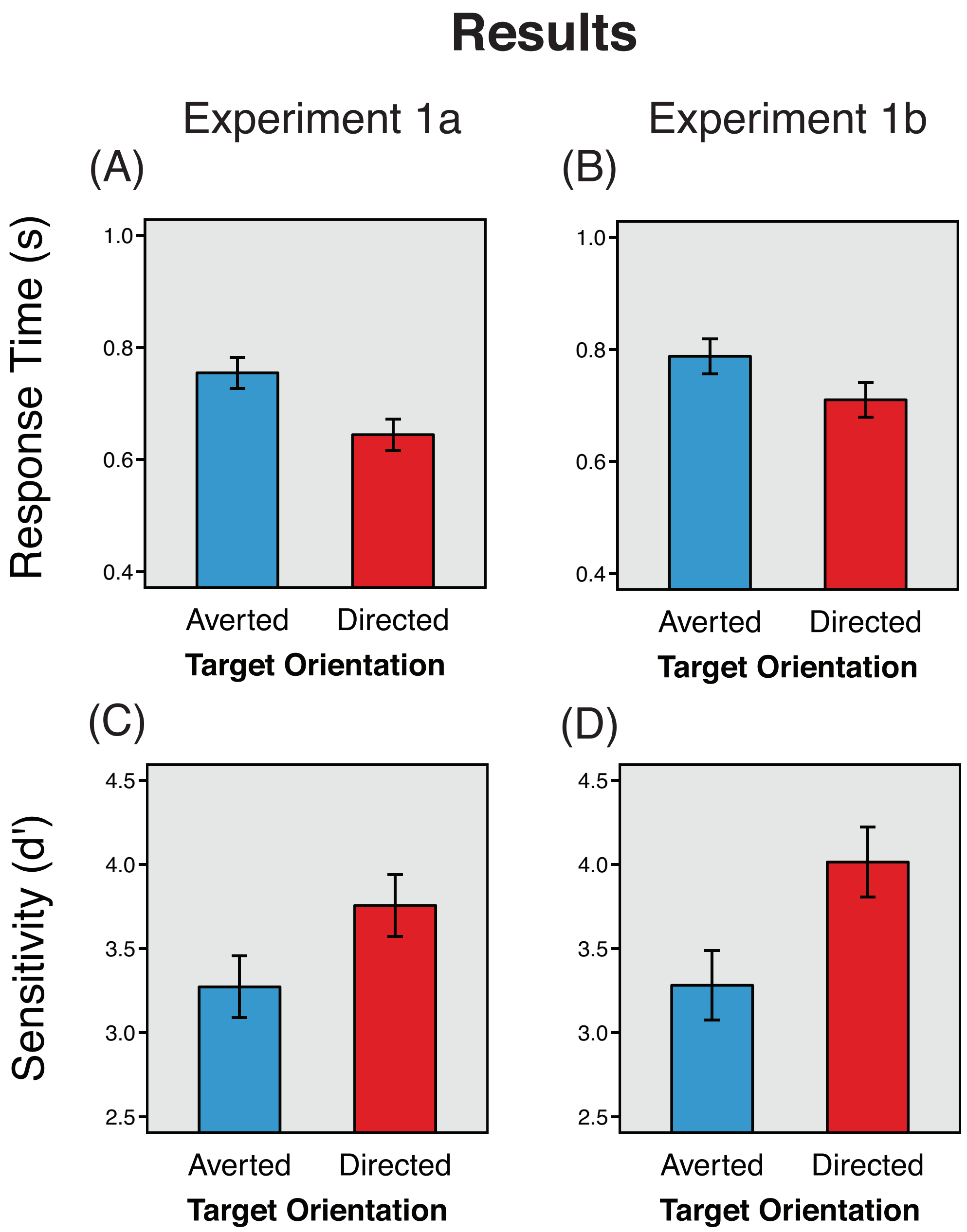

Figure 3.2: Results from Experiments 1a and 1b. (a) Average response times for Find Averted versus Find Directed displays in Experiment 1a. (b) Average response times for Find Averted versus Find Directed displays in Experiment 1b. (c) Sensitivity (measured as d') for Find Averted versus Find Directed displays in Experiment 1a. (d) Sensitivity (measured as d') for Find Averted versus Find Directed displays in Experiment 1b. Error bars reflect 95 % confidence intervals, subtracting out the shared variance.

3.5 Experiment 2a: Symmetry Control

The results of Experiment 1 suggest that *stare-in-the-crowd* effects might occur even for eyeless stimuli that may nevertheless convey a sense of directed attention or intention. However, our directed and averted stimuli also differed in another way that did not signal agency: might the directed cones have been more readily detected simply because they were more symmetric? To find out, we replicated Experiment 1, but now the cones were replaced with stimuli that retained the same differential symmetry but did not convey any directedness in the first place (as depicted in Fig. 3.3a).

3.5.1 Method

This experiment was identical to Experiments 1a and 1b, except as noted here. Ten members of the Yale community (M_{age} = 19.1) participated in exchange for course credit (with this sample size chosen to exactly match that of Experiments 1a and 1b). Instead of cones, each stimulus consisted of a horizontally oriented dark-gray rectangle (2.94° $\times$ 1.94°, with a 0.04° white border segmenting it from the darker background), with an inset vertical bar (0.78° wide). The bar's center was colored white, with a continuous gradient from light to dark on either side of this vertical line. (Each bar appeared as a gradient in this way in order to match the sort of gradual shading and three-dimensional appearance employed in the cones from Experiment 1. Indeed, the extreme shades of this gradient were chosen to exactly the match the lightest and darkest regions of the cones.) For symmetric stimuli (Fig. 3.3a, bottom), the vertical bar was horizontally centered in the rectangle; for asymmetric stimuli (Fig. 3.3a, top), the bar was shifted (either to the left or right) by 0.91°. These displays thus consisted of clearly symmetric versus asymmetric patterns, but without any sense of directedness as in Experiment 1. (Due to the gradient of each bar, these stimuli also looked three-dimensional, but instead of a cone that was pointing in a specific direction, each bar simply appeared as a kind of vertical pole viewed through an aperture.)

3.5.2 Results and Discussion

We analyzed an average of 141.00 nonexcluded trials per observer. As depicted in Fig. 3.3b, responses were faster for Find Asymmetric (compared to Find Symmetric) displays (0.76 vs. 0.93 s, $t(9) = 6.29$, $p < .001$, $d = 1.87$). Critically, this effect was significantly different (and of course was in the *opposite* direction) from those of both Experiment 1a (0.11 vs. −0.17 s, $t(18) = 7.14$, $p < .001$, $d = 3.19$) and Experiment 1b (0.08 vs. −0.17 s, $t(18) = 6.00$, $p < .001$, $d = 2.68$). Sensitivity (again computed as *d'*) was also higher for Find Asymmetric (compared to Find Symmetric) displays (3.89 vs. 2.34, $t(9) = 6.47$, $p < .001$, $d = 2.32$). And this effect was also significantly different (and again in the opposite direction) from those of both Experiment 1a (−0.48 vs. 1.54, $t(18) = 6.68$, $p < .001$, $d = 2.99$) and Experiment 1b (−0.73 vs. 1.54, $t(18) = 7.13$, $p < .001$, $d = 3.19$).

These results clearly indicate that the search advantage for directed (vs. averted) cones in Experiment 1 was unlikely to be due to their differential (a)symmetry. Indeed, if anything the current results suggest that Experiment 1 may have *underestimated* the magnitude of the stare-in-the-crowd effect with cones, since the brute effect of symmetry went so strongly in the opposite direction.

3.6 Experiment 2b: Direct Replication

Given the importance of direct replications, we reran the experiment on another group of 10 observers ($M_{age} = 19.1$), analyzing an average of 143.30 nonexcluded trials per observer. As depicted in Fig. 3.3c, responses were again faster for Find Asymmetric (compared to Find Symmetric) displays (0.73 vs. 0.95 s, $t(9) = 5.21$, $p < .001$, $d = 2.04$). Critically, this effect was significantly different (and again in the *opposite* direction) from those of both Experiment 1a (0.11 vs. −0.23 s, $t(18) = 6.47$, $p < .001$, $d = 2.89$) and Experiment 1b (0.08 vs. −0.23 s, $t(18) = 5.66$, $p < .001$, $d = 2.53$). Sensitivity (again computed as d') was also again higher for Find Asymmetric (compared to Find Symmetric) displays (3.96 vs. 2.37, $t(9) = 12.26$, $p < .001$, $d = 3.01$). And this effect was also significantly different (and again in the opposite direction) from those of both Experiment 1a (−0.48 vs. 1.59, $t(18) = 9.10$, $p < .001$, $d = 4.07$) and Experiment 1b (−0.73 vs. 1.59, $t(18) = 9.35$, $p < .001$, $d = 4.18$).

3.7 Experiment 3a: Contrast Control

The Directed and Averted stimuli employed in Experiment 1 differed not just in their symmetry, but also in their contrast; might the directed cones have been more readily detected simply because of their higher contrast? To find out, we replicated Experiment 1, but now the cones were replaced with stimuli that had clearly differential contrast but did not convey any directedness in the first place, as depicted in Fig. 3.3d.

3.7.1 Method

This experiment was identical to Experiments 1a and 1b, except as noted here. Ten members of the Yale community ($M_{age} = 21.5$) participated (with this sample size chosen to exactly match that of Experiments 1a, 1b, 2a, and 2b). Instead of cones, each stimulus consisted of a pinwheel (1.94° × 1.94°), subdivided into four light-gray (base width: 0.58°) and four dark-gray (base width: 0.86°) alternating wedges (Fig. 3.3d, bottom). For lower contrast stimuli, a global 0.06° Gaussian blur filter was applied (Fig. 3.3d, top). These displays thus consisted

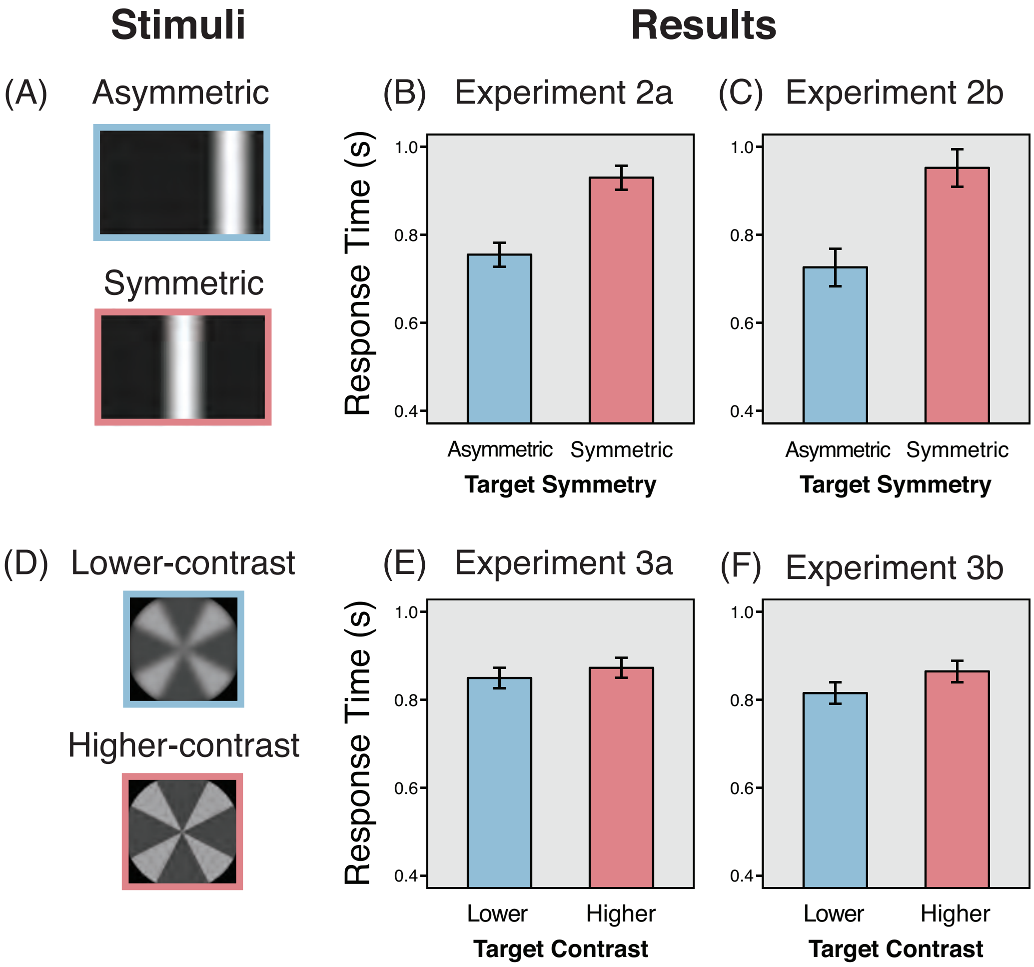

Figure 3.3: Stimuli and results from Experiments 2a, 2b, 3a, and 3b. (a) Asymmetric and symmetric targets employed in Experiments 2a and 2b . (b) Average response times for Find Asymmetric versus Find Symmetric displays in Experiment 2a. (c) Average response times for Find Asymmetric versus Find Symmetric displays in Experiment 2b. (d) Lower- and Higher contrast targets employed in Experiments 3a and 3b. (e) Average response times for Find Lower- versus Find Higher contrast displays in Experiment 3a. (f) Average response times for Find Lower- versus Find Higher contrast displays in Experiment 3b. Error bars reflect 95 % confidence intervals, subtracting out the shared variance.

of clearly higher- versus lower contrast patterns, but without any sense of directedness as in Experiment 1.

3.7.2 Results and Discussion

We analyzed an average of 131.60 nonexcluded trials per observer. As depicted in Fig. 3.3e, response times for Find Higher Contrast versus Find Lower Contrast displays did not differ (0.87 vs. 0.85 s, $t(9) = 0.99$, $p = .349$, $d = 0.12$). Critically, this effect was significantly different from those of both Experiment 1a (0.11 vs. −0.02 s, $t(18) = 3.60$, $p = .002$, $d = 1.61$) and Experiment 1b (0.08 vs. −0.02 s, $t(18) = 2.56$, $p = .020$, $d = 1.15$). Sensitivity (again computed as *d'*) also did not differ between Find Higher Contrast and Find Lower Contrast displays (2.19 vs. 2.60, $t(9) = 1.73$, $p = .117$, $d = 0.33$)—and of course this effect was numerically in the *opposite* direction from that in Experiments 1a and 1b. And this effect was also significantly different from those of both Experiment 1a (−0.48 vs. 0.42, $t(18) = 2.95$, $p = .009$, $d = 1.32$) and Experiment 1b (−0.73 vs. 0.42, $t(18) = 3.59$, $p = .002$, $d = 1.60$).

These results clearly indicate that the search advantage for directed (vs. averted) cones in Experiment 1 was unlikely to be due to differential contrast.

3.8 Experiment 3b: Direct Replication

Given the importance of direct replications, we reran the experiment on another group of 10 observers ($M_{age} = 21.4$), analyzing an average of 135.30 nonexcluded trials per observer. As depicted in Fig. 3.3f, responses for Find Higher Contrast versus Find Lower Contrast displays did not differ (0.86 vs. 0.81 s, $t(9) = 1.97$, $p = .080$, $d = 0.26$)—and this marginal effect was in the *opposite* direction from that in Experiment 1 (consistent with the idea that if anything, Experiment 1 may have underestimated the effect of perceived intentionality). Critically, this effect was significantly different from those of both Experiment 1a (0.11 vs. 0.05 s, $t(18) = 4.19$, $p = .001$, $d = 1.87$) and Experiment 1b (0.08 vs. 0.05 s, $t(18) = 3.15$, $p = .006$, $d = 1.41$). Sensitivity also did not differ between Find Higher Contrast and Find Lower Contrast displays (2.67 vs. 2.73, $t(9) = 0.46$, $p = .657$, $d = 0.06$). And this effect was also significantly different from those of both Experiment 1a (−0.48 vs. 0.06, $t(18) = 2.34$, $p = .032$, $d = 1.05$) and Experiment 1b (−0.73 vs. 0.06, $t(18) = 3.14$, $p = .007$, $d = 1.40$).

3.9 GENERAL DISCUSSION

This study makes a simple but important point: the stare-in-the-crowd effect obtains not only with faces and eyes (as has been demonstrated in dozens of previous experiments) but also with directed cone shapes that appear to be *facing* you. (And since the effect vanished—or even reversed—with the control stimuli in both Experiments 2 and 3, it seems unlikely that this effect could merely reflect an influence of differential symmetry or contrast.) These cone stimuli are radically different from faces at the level of superficial visual features—and of course they are entirely eye-less, by design. But they nevertheless have one critical property in common: because of their coordinated orientations, the cones may also convey a sense of directed attention and intention. Accordingly, we conclude that the advantage of direct stares in visual search is not specific to eyes after all: it can also arise with *staring* shapes[I]. And this kind of empirical pattern might itself generalize even beyond visual search and the stare-in-the-crowd effect. Indeed, we may have been incorrectly limiting the scope of a much wider

---

[I] Of course, it is also possible that these results could reflect an entirely new and different effect, with no connection to gaze, but this would require positing an entirely new and unmotivated mechanism (which seems less parsimonious)—whereas the framework we are suggesting posits that the current effect and the traditional 'stare-in-the-crowd' effect with eyes are one and the same phenomenon.

class of effects to only *eye contact*, when in fact they also generalize to other types of *intentional* cues. Based on the present results, for example, we suspect that effects of direct eye gaze on cognitive control (Conty et al., 2010) and working memory (Wang & Apperly, 2017) might similarly also arise with directed (but eye-less) shapes (for an example, see Colombatto et al., 2019). In other words, all of these kinds of effects of direct eye contact might instead reflect a more general phenomenon of *mind contact*.

4

Gaze deflection
reveals how gaze cueing is tuned
to extract the mind behind the eyes

4.1 Abstract

Suppose you are surreptitiously looking at someone, and then when they catch you staring at them, you immediately turn away. This is a social phenomenon that almost everyone experiences occasionally. In such experiences—which we will call gaze deflection—the "deflected" gaze is not directed at anything in particular but simply away from the other person. As such, this is a rare instance where we may turn to look in a direction without intending to look there specifically. Here we show that gaze cues are markedly less effective at orienting an observer's attention when they are seen as deflected in this way—even control-

ling for low-level visual properties. We conclude that gaze cueing is a sophisticated mental phenomenon: It is not merely driven by perceived eye or head motions but is rather well tuned to extract the "mind" behind the eyes.

4.2 Significance

We report an empirical study of gaze deflection—a common experience in which you turn to look in a different direction when someone "catches" you staring at them. We show that gaze cueing (the automatic orienting of attention to locations at which others are looking) is far weaker for such displays, even when the actual eye and head movements are identical to more typical intentional gazes. This demonstrates how gaze cueing is driven by the perception of minds, not eyes, and it serves as a case study of both how social dynamics can shape visual attention in a sophisticated manner and how vision science can contribute to our understanding of common social phenomena.

4.3 Introduction

One of the most important events we perceive in our daily lives is when a nearby agent shifts their attention, e.g., turning suddenly to look in a different direction. Indeed, our visual system is especially sensitive to where others are looking, as demonstrated by many previous studies of gaze shifting (for a review see Emery, 2000), and these events are so salient that we have an automatic tendency to look in the direction that others are looking (Milgram, Bickman, & Berkowitz, 1969). This gives rise to the phenomenon of gaze cueing: In a display with two potential target locations flanking a face, for example, observers are faster and more accurate at identifying targets that appear where the face is looking (e.g., Driver et al., 1999;

Friesen & Kingstone, 1998; for a review, see Frischen, Bayliss, & Tipper, 2007). This sort of gaze cueing is triggered not just when viewing eyes but also when viewing simple head turns (Langton & Bruce, 2000; see also Langton, Watt, & Bruce, 2000).

Why are such gaze shifts so powerful? They might be driven simply by the salient motions of the eyes and heads themselves. But another possibility is that they are driven by the higher-level perception that an agent has shifted their attention or intentions. Exploring these possibilities requires a stimulus in which these factors diverge, which may seem unusual; after all, we usually look toward the objects that are the focus of our intentions (Krajbich, Armel, & Rangel, 2010). But there is one relatively common (though previously unstudied) social phenomenon in which a gaze shift may not actually signal an intention to look at the second location. This occurs in what we will call gaze deflection—when you are surreptitiously looking at someone but then suddenly look away (perhaps toward a second person) when the first person catches you staring at them. Here the intention is not to look at the second person, but only away from the first person.

Do such "deflected" gazes still drive gaze cueing? In five experiments (including direct replications), we showed each observer an animation with three actors[1] either exhibiting gaze deflection (deflection animations) or performing identical movements, except now temporally reordered, such that impressions of gaze deflection were eliminated and all gaze shifts were seen as intentionally directed at their new locations (control animations). In Experiment 1a, each animation (depicted in Fig. 4.1 and also online at http://www.yale.edu/perception/gaze-deflection/) began with a central person (A) turning to look at the rightmost person (B; the "first" gaze, seen as directed). In deflection animations, B turned her head to face A, who then (exhibiting gaze deflection) immediately turned to look in the other direction, thus fac-

---

[1] In fact, the people in the videos were the paper's authors—but for reasons of agreed- upon differences in photogenic fitness, one of the authors was included twice, and one was eliminated altogether.

ing a third person (C; the "second" gaze, seen as deflected). In control animations, shortly after turning to look at B, A instead spontaneously (i.e., without B "catching" her staring) turned to look toward C (the second gaze, now seen as directed). Only then did B turn her head toward A. To measure how observers' attention varied in response to the deflection vs. control animations, we presented a single target letter along the direction of A's gaze during either the first gaze (early targets; depicted in Fig. 4.2a) or second gaze (late targets; depicted in Fig. 4.2b). This same design was then employed in Experiment 1b (a direct replication of Experiment 1a).

Next, we ruled out two classes of potential confounding factors, pertaining to temporal differences (Experiment 2) and spatial differences (Experiments 3a and 3b) in the animations employed in the original experiments. In Experiment 2, we explored the role of temporal factors: Whereas the deflection vs. control animations in Experiments 1a and 1b featured different numbers of head turns (and differential delays) before the late target was presented, these temporal factors were now equated (as depicted in Supplementary Fig. 4.3). This experiment also served as a conceptual replication of Experiments 1a and 1b since they featured different videos, now of actors facing toward the camera so that their eyes were fully visible (as in Fig. 4.2d and 4.2e).

Finally, we explored the role of spatial factors: In Experiment 2, deflection animations ended with both actors looking toward the target location (as in Fig. 4.2d), while control animations ended with one of the actors looking forward (as in Fig. 4.2e). To ensure that these differing spatial configurations could not explain the observed differences between deflected and control animations, Experiment 3a (and Experiment 3b, its direct replication) retained these final tableaus from Experiment 2 but eliminated the preceding motions which led to the perception of deflected vs. directed gazes in the first place (as depicted in Supplementary Fig. 4.4).

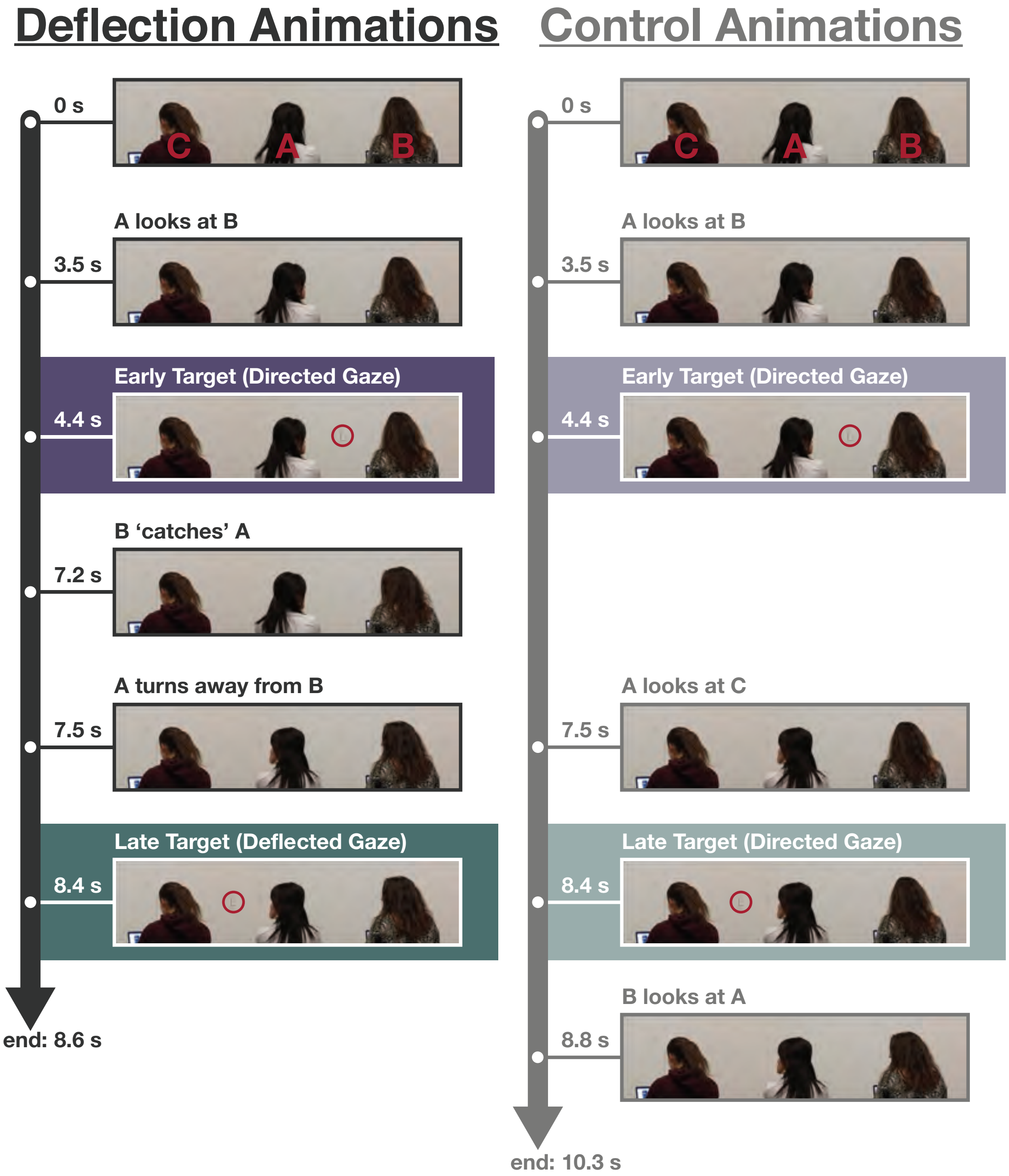

Figure 4.1: A schematic depiction of the animations observers viewed in Experiments 1a and 1b.

4.4 RESULTS

The average discrimination accuracy for early and late targets in Experiment 1a is depicted separately for deflection and control animations in Fig. 4.2c. Inspection of this figure suggests two clear patterns of results: 1) In the deflection animations, letter discrimination accuracy was higher for first gazes (seen as directed) compared to second gazes (seen as deflected), but 2) this bias was not present in the control animations (when both gazes were directed). These impressions were confirmed by the following analyses. The proportions of correct responses for early and late targets were compared using a two-proportion z test in deflection and control animations, respectively. There was a significant difference between first and second gazes in the deflection animations (87.0 vs. 71.0 %; $z = 2.78$, $p = 0.005$, Cohen's $h = 0.40$) but not in control animations (74.0 vs. 80.0 %; $z = 1.01$, $p = 0.313$, $h = 0.14$). And the difference between these differences (i.e., the interaction effect) was also highly reliable ($z = 2.66$, $p = 0.008$).

Thus, the gaze cueing effect is greatly reduced when the gaze is deflected, even when the actual head motion is identical. These effects were directly replicated in Experiment 1b: there was a significant difference between early and late targets in deflection animations (86.0 vs. 70.0 %; $z = 2.73$, $p = 0.006$, $h = 0.39$) but not in control animations (74.0 vs. 76.0 %; $z = 0.33$, $p = 0.744$, $h = 0.05$), and the difference between these differences was also significant ($z = 2.12$, $p = 0.034$). And in Experiment 2, gaze cueing was once again greatly reduced when the gaze was deflected (61.0 vs. 71.0 %; $z = 2.11$, $p = 0.035$, $h = 0.21$; see Fig. 4.2f)—despite the identical timing of head turns, even with fully visible eyes.

Experiment 3a retained only the final tableaus from Experiment 2. Now, with the elimination of the preceding head turns (that yielded impressions of gaze deflection), there was no difference between the deflection-frame and control-frame conditions. Indeed, if anything, there was a trend in the opposite direction: Accuracy was greater with final tableaus from deflection animations compared to those from control animations (54.0 vs. 48.0 %; z = 1.20, p = 0.230, h = 0.12; see Fig. 4.2f). And crucially, there was a reliable interaction with target accuracy from Experiment 2 (z = 2.32, p = 0.020), thus demonstrating that the final spatial configurations alone cannot be responsible for the gaze deflection effect. These effects were directly replicated in Experiment 3b: accuracy was again trending in the opposite direction, with better performance for final tableaus from deflection compared to control animations (55.0 vs. 45.5 %; z = 1.90, p = 0.057, h = 0.19), and the interaction again revealed that this was different from the gaze deflection effect observed in Experiment 2 (z = 2.83, p = 0.005).

4.5 Discussion

The current study exploits the phenomenon that we have called gaze deflection. This is a familiar (perhaps all too familiar) social phenomenon from everyday life, whereas most studies of gaze cueing use either static images of isolated faces or short video clips in which eye movements are divorced from their context (such that the agents in most such experiments are not actually looking at anything; see Langton, 2009). The results were clear and powerful: even when tested in only a single trial per observer, eye and head movements were much less effective at cueing attention when they were seen as deflected and thus dissociated from the actual direction of intention.

These effects seemed to reflect the social significance of gaze deflection, rather than any lower-level properties. In particular, they did not reflect differences in the timing of head turns since these were equated in Experiment 2. And they also cannot be explained by traditional gaze cueing mechanisms to the differing spatial configurations. When multiple people turn to look in the same direction (as in Fig. 4.2d), gaze cueing is typically amplified (e.g., Gallup et al., 2012), and such "pooling" effects are particularly strong in the context of actual

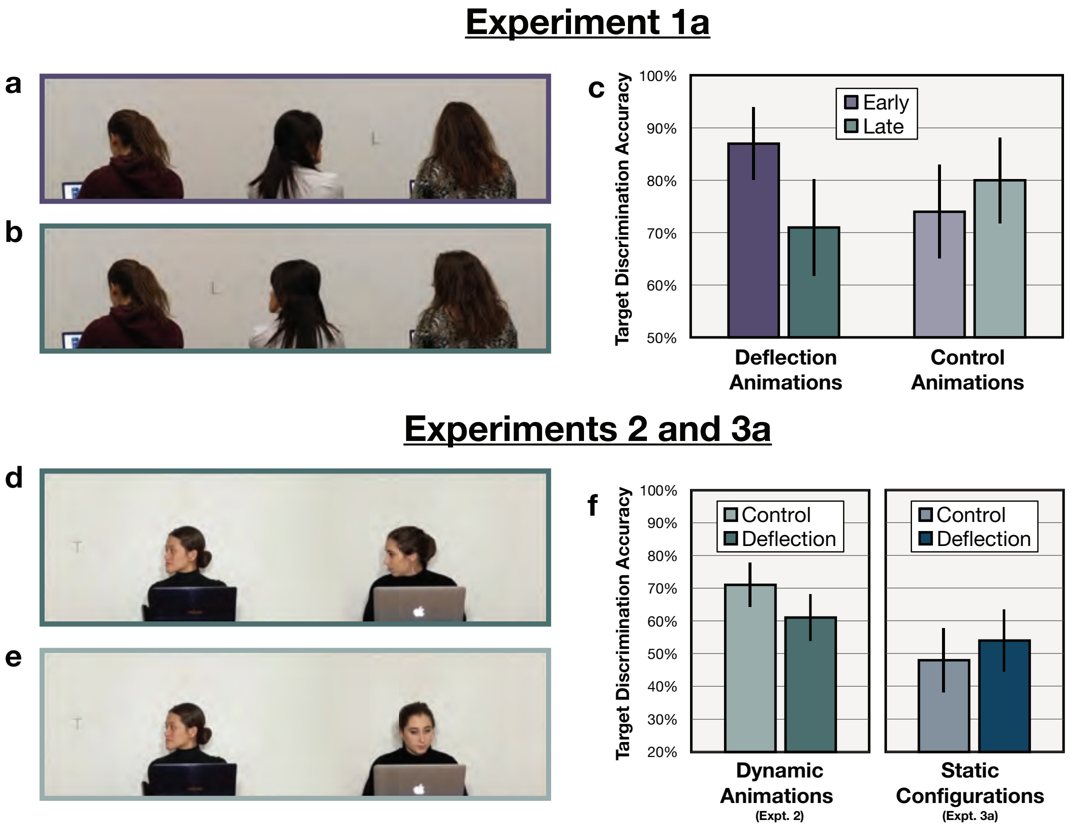

Figure 4.2: Stimuli and results from Experiments 1a, 2, and 3a. (a and b) Examples of the stimuli used in the letter discrimination task from Experiments 1a and 1b, including the early target (a) and late target (b). (c) Average accuracy in the letter discrimination task for early vs. late targets in Experiment 1a. Accuracy was impaired for late targets (relative to early targets) in deflection animations but not in control animations (where both gazes were seen as directed since there was no gaze deflection). (d and e) Examples of the stimuli used in the letter identification task from Experiment 2 in both the deflection animation (d) and the control animation (e). (f) Average accuracy in the letter identification task for deflection vs. control animations in Experiment 2 and for deflection vs. control configurations in Experiment 3a. Accuracy was impaired for targets (relative to control animations) in deflection animations, but only when presented as dynamic animations (Experiment 2) and not when presented as static frames (thus eliminating impressions of gaze deflection; Experiment 3a). Error bars indicate 95 % CIs.

head turns as we used in the present studies (as opposed to mere eye movements; e.g., Florey et al., 2016). This sort of group-wide gaze cueing remains powerful even when people are looking directly to the right or left (Sun et al., 2017), and attention is not cued in such configurations to the space between multiple people who are gazing in the same direction (Vestner et al., 2019). Accordingly, we also demonstrated directly (in Experiments 3a and 3b) that such spatial configurations do not yield such differences in the absence of the head turns and eye movements that give rise to the perception of gaze deflection.

The phenomenon of gaze deflection, with its dissociation between the perceived direction of gaze and the perceived direction of intention, provides unique insights into recent debates on the relative contribution of visual cues and mental states to social attention (for a recent review, see Capozzi & Ristic, 2020). It has long been assumed that gaze cueing is driven by the visual cue of eye gaze alone (Friesen & Kingstone, 1998; Kingstone et al., 2019). Building on other recent work uncovering humans' remarkable ability to construct rich models of others' attentional states (e.g., Guterstam et al., 2019), the current results use a familiar social phenomenon to directly demonstrate that cueing of attention is especially tuned to the perceived attentional states of others and less so to brute visual cues. Attention, in this sense, seems tuned not to follow the eyes, but rather to follow the mind behind the gaze.

4.6 Methods and Materials

4.6.1 Participants

For each experiment, 400 observers were recruited through Amazon Mechanical Turk (MTurk; Experiment 1a: 244 females, M_{age} [the participants' mean age] = 36.96; Experiment 1b: 242 females, M_{age} = 35.77; Experiment 2: 203 females, M_{age} = 34.40; Experiment 3a: 187 females,

M_{age} = 36.21; Experiment 3b: 194 females, M_{age} = 34.67), and each completed a single trial in a 2- to 5-min session in exchange for monetary compensation. (For a discussion of this pool's nature and reliability, see Crump, McDonnell, & Gureckis, 2013). All observers were in the United States, had an MTurk task approval rate of at least 80 %, and had previously completed at least 50 MTurk tasks.) This sample size was determined arbitrarily before data collection began and was fixed to be identical in each of the three experiments reported here. All experimental methods and procedures were approved by the Yale University Institutional Review Board, and all observers confirmed that they had read and understood a consent form outlining their risks, benefits, compensation, and confidentiality and that they agreed to participate in the experiment.

4.6.2 Apparatus

After agreeing to participate, observers were redirected to a website where stimulus presentation and data collection were controlled via custom software written in HTML, CSS, JavaScript, and PHP. (Since the experiment was rendered on observers' own web browsers, viewing distance, screen size, and display resolutions could vary dramatically, so we report stimulus dimensions below using pixel [px] values.)

4.6.3 Stimuli and Design

Experiments 1a and 1b

As depicted in the sample screenshots in Fig. 4.1, observers viewed an animation (1,202 × 297 px) centered in their browser window including a gray (hexadecimal color code #605D5D) 3 px frame on a dark gray (#404040) background. Three people were viewed from behind, on a background wall (approximately #CFCBC4). The three people were sitting in front of laptops and typing sounds played throughout the animation. The people initially looked straight ahead, with the timings of the movements described below reported with respect to the beginning of the animation.

In the deflection animations, the central person turned her head (at 3.5 s) toward the rightmost person (the first gaze, seen as directed) and then seemed to stare at her. At 7.2 s, the rightmost person turned her head to face the middle person, who then (exhibiting gaze deflection) immediately (at 7.5 s) turned to look in the other direction, thus facing the leftmost person (the second gaze, seen as deflected). (The leftmost person looked straight ahead throughout the animation.) The final tableau was then visible for an additional 1.1 s (i.e., until 8.6 s), at which point it disappeared. In the control animations, the central person again turned her head (at 3.5 s) toward the rightmost person (the first gaze, again seen as directed) and then seemed to stare at her. At 7.5 s (without having been caught), the central person then turned to look in the other direction, thus facing the leftmost person (the second gaze, also now seen as directed). Only after this (at 8.8 s) did the rightmost person turn her head toward the central person. The final tableau was then visible for an additional 1.5 s (i.e., until 10.3 s), at which point it disappeared. (Once again, the leftmost person looked straight ahead throughout the animation.)

Each observer viewed a target letter presented for 0.13 s on the background between the people (roughly in line with their eyes) while the animation was playing. This target was a gray (#9C9892) "T" or "L" (presented in Helvetica, roughly 20 × 30 px). Targets presented during first gazes (early targets; depicted in Fig. 4.2a) were presented between the middle and rightmost people (centered at 788 px from the image's left border) along the direction of gaze (centered at 143 px from the image's top border) 0.1 s after the middle person finished turning her head toward the rightmost person (at 4.4 s). Targets presented during second gazes (late targets; depicted in Fig. 4.2b) were presented between the middle and leftmost people (centered at 408 px from the image's left border) along the direction of gaze (centered at 143 px from the image's top border) 0.1 s after the middle person finished turning her head toward the leftmost person (at 8.4 s).

In the actual animations that observers viewed, the identities of the leftmost and rightmost people were counterbalanced, using the identical stimuli. In fact, since the leftmost person never turned her head, only two initial movies were filmed, but the leftmost person in each movie was the first static frame of the rightmost person from the other movie. (Given the uniformly lit wall in the background, this frame was added into the animation without any obvious segmentation cue, such that it appeared to be an animation of three separate people, as depicted in Fig. 4.1.) The two resulting animations were qualitatively identical, but because they were constructed from two separately filmed movies, their timing was slightly different. In particular, compared to the timing of the first pair of animations (as described above), the second movie's key events occurred at the following time stamps: 1) The middle person turned to the right at 3.7 s. 2) In the deflection animations, the rightmost person then turned to the left at 7.2 s. 3) In the deflection animations, the middle person turned to the left at 7.6 s. 4) In the control animations, the middle person turned to the left at 7.6 s. 5) In the control animations, the rightmost person looked to the left at 9.0 s. 6) Targets presented during first gazes appeared at 5.4 s. 7) Targets presented during second gazes appeared at 8.4 s.

The design described above resulted in a total of 16 animations: 2 target timings (early/late) × 2 target identities (L/T) × 2 orders of head movements (deflection/control) × 2 identities for the rightmost vs. leftmost people, and each was viewed by 25 unique observers.

EXPERIMENT 2

Observers viewed a silent animation (1,000 × 298 px) including a gray (#5F5D5B) 6 px frame and featuring two people viewed from the front on a background wall (approximately #DFDFD7).

In the deflection animation, the left person turned her head (at 2.0 s) toward the right person and then seemed to stare at her. At 4.0 s, the right (i.e., stared-at) person turned her head to face the left person, who then (exhibiting gaze deflection) immediately (at 4.8 s) turned to look in the other direction. The final tableau was then visible for an additional 0.7 s (i.e., until 6.0 s), at which point it disappeared. In the control animation, the right person was facing to her left in the beginning, and at 2.0 s she turned to face straight ahead. At 4.0 s, the left person turned her head toward the right person to stare at her and then immediately (at 4.8 s) turned to look in the other direction. (Since the right person was facing her laptop during these movements, this shift now appeared to be intentional rather than deflected.) The final tableau was then visible for an additional 0.7 s (i.e., until 6.0 s), at which point it disappeared.

A target letter was presented to the left of the left person (centered at 130 px from the image's left border) 0.1 s after the left person finished turning her head toward her right (and the observer's left) along the direction of gaze (centered at 283 px from the image's top border). This target (a gray #B1B0A7 T, presented in Helvetica, roughly 38 × 46 px) gradually faded in over the course of 0.20 s, remained visible for 0.10 s, and gradually faded out for another 0.20 s. There were thus two animations corresponding to two orders of eye/head motions (deflection/control), and each was viewed by 200 unique observers.

EXPERIMENTS 3A AND 3B

These experiments were identical to Experiment 2, except as noted here. Observers viewed the final 0.7 s of the animations (which consisted of only static frames) from Experiment 2 (cropped to hide the laptop logos, 1,000 × 268 px). In deflection frames, both people were thus facing to the left (as in Fig. 4.2d), and in control frames, the left person was facing to the left, while the right person was facing ahead (as in Fig. 4.2e). The target letter (#000000) was again presented to the left of the left person and began fading in as the animation began. The animation ended 0.2 s after the target faded out, at 0.7 s. There were thus two animations corresponding to two head directions (deflection/control frames), and for half of the observers (counterbalanced across conditions), the videos were horizontally flipped, for a total of four animations, each viewed by 100 unique observers.

4.6.4 PROCEDURE

Each observer was instructed to watch a single animation as closely as possible, as it would be displayed only once. Observers viewed the animation (which started playing automatically after 0.5 s in Experiments 1a and 1b, upon a keypress in Experiment 2, and upon a keypress and after a 1 s "Get Ready" message in Experiments 3a and 3b). In Experiments 1a and 1b, immediately after the animation ended (and disappeared), observers were asked three questions (only one of which was visible at a time): 1) whether they had seen a letter appear during the animation, 2) whether it was a "T" or an "L" (and to guess if they did not know), and 3) how confident they were in their response (on a scale of 1 to 7, with 1 labeled "Not at all" and 7 labeled "Entirely"). In Experiment 2, observers were asked two questions (only one of which was visible at a time): 1) whether they had seen a letter appear during the animation and 2) which letter they saw (A to Z; and to guess if they did not know). In Experiments 3a and 3b, observers were asked only one question: which letter they saw (A to Z; and to guess if they did not know). In all experiments, they then also answered questions that allowed us to exclude (with replacement) observers who guessed the purpose of the experiment (e.g., mentioning gaze following; $n = 27, 10, 25, 15$, and 7 in Experiments 1a, 1b, 2, 3a, and 3b respectively), who interrupted the experiment ($n = 16, 9, 21, 55, 54$), who did not view the video "in full view" ($n = 95, 49, 9, 11, 11$), who reported past participation in a similar study ($n = 4, 8, 24, 46, 45$), who encountered any problems ($n = 19, 7, 4, 3, 2$), or who failed to answer our questions sensibly ($n = 7, 2, 24, 49, 32$; e.g., responding to our question about the experiment's purpose by writing "i cant see"). In Experiments 2, 3a, and 3b, we also removed observers who entered anything other than a single letter in response to the letter identification question (e.g., "jjhhgkjk"; $n = 2, 17, 9$). The resulting unique excluded observers (some of whom triggered multiple criteria; $n = 142, 70, 79, 101, 96$) were replaced without us ever analyzing their data. (The relatively high exclusion rate for observers who reported not watching the video in full view in Experiments 1a and 1b may be due to observers misunderstanding our poorly worded question as involving whether the people in the videos—and not the videos themselves—were in full view. In fact, the video depicted only the upper bodies of the people, as in Fig. 4.1. When this question was replaced by directly measuring the size of observers' browser windows in Experiments 2, 3a, and 3b and comparing it to the size of the animation, only 9, 11, and 11 observers were excluded.)

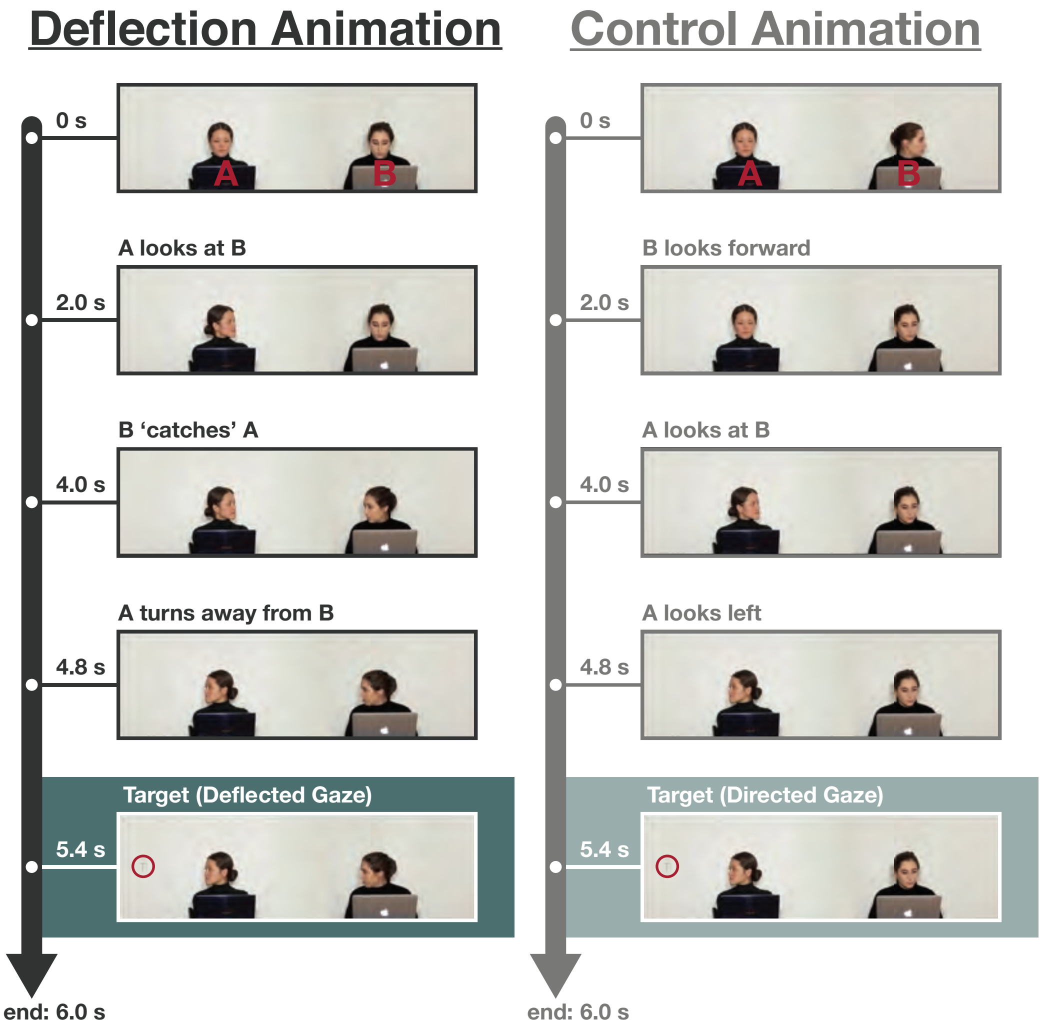

Figure 4.3: A schematic depiction of the animations observers viewed in Experiment 2.

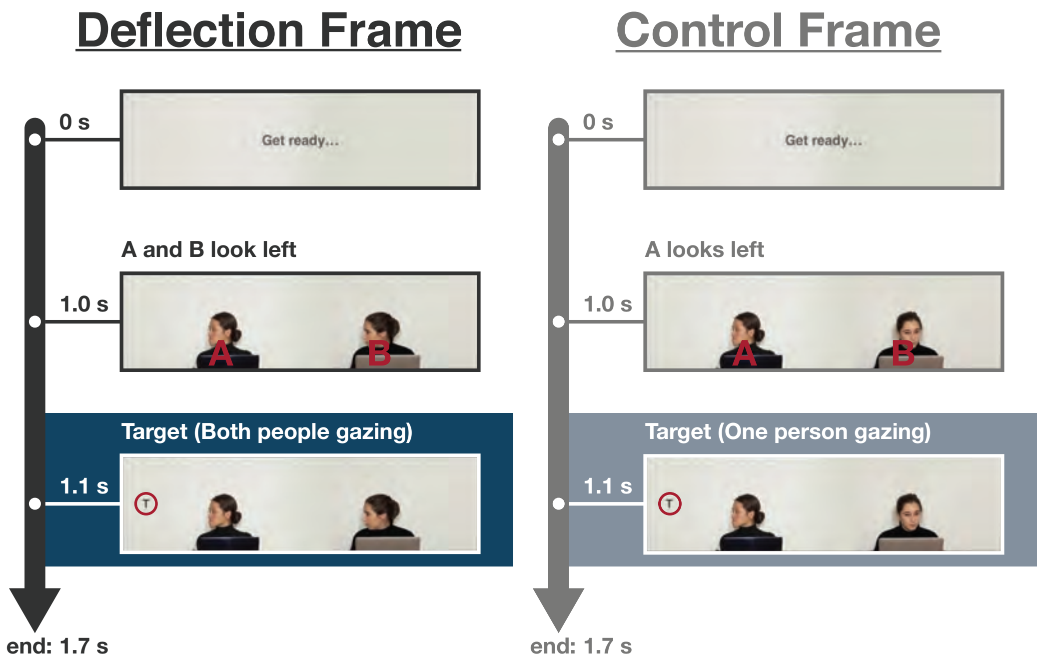

Figure 4.4: A schematic depiction of the frames observers viewed in Experiments 3a and 3b.

5

Unconscious pupillometry:

An effect of 'attentional contagion' in the absence of visual awareness

5.1 Abstract

When looking at other people, we can readily tell how attentive (or distracted) they are. Some cues to this are fairly obvious (as when someone stares intensely at you), but others seem more subtle. For example, increased cognitive load or emotional arousal causes one's pupils to dilate. This phenomenon is frequently employed as a physiological *measure* of arousal, in studies of pupillometry. Here, in contrast, we employ it as a *stimulus* for social perception. Might the human visual system be naturally and automatically engaging in 'unconscious pupillometry'? Here we demonstrate that faces rendered invisible

(through continuous flash suppression) enter awareness faster when their pupils are dilated. This cannot be explained by appeal to differential contrast, differential attractiveness, or spatial attentional biases, and the effect vanishes when the identical stimuli are presented in socially meaningless ways (e.g., as shirt buttons or facial moles). These results demonstrate that pupil dilation is prioritized in visual processing even outside the focus of conscious awareness, in a form of unconscious 'attentional contagion'.

5.2 INTRODUCTION

A critical task for vision is determining what we should attend to, in what might otherwise be an overwhelming stream of sensory input. Often, of course, we simply make voluntary decisions about what to focus on. But attention is also more automatically attracted to certain categories of stimuli—especially other people (e.g., New et al., 2007; Ro et al., 2001). This may be an adaptive bias, insofar as other people are more likely than most other stimuli (e.g., trees or clouds) to act in a way that may directly impact our fitness. But this is not equally true for all people: those who are actively attending (especially to us) may be much more likely to immediately influence our welfare, compared to people who are inattentive—and in fact, people spend a rather amazing amount of time being distracted (Killingsworth & Gilbert, 2010) or focusing internally rather than externally (Chun et al., 2011).

So how can we tell whether someone is attentive or distracted? Some of the cues seem obvious—as when someone turns to look in a particular direction (e.g., Milgram et al., 1969), is looking directly at us (von Grünau & Anston, 1995), or stops blinking or moving their eyes (e.g., Reichle et al., 2010; Smilek et al., 2010). But other cues seem more subtle. Perhaps the best example of this is *pupil size*: our pupils dilate when we are attentionally engaged—e.g., as the result of heightened interest (Hess & Polt, 1960), increased cognitive load (Kahneman

& Beatty, 1966), emotional arousal (Bradshaw, 1967), or uncertainty (Lavín et al., 2014). Indeed, pupils dilate obligatorily upon excitation of the nervous system (Applegate et al., 1983; Reimer et al., 2016), and even unbeknownst to the subject (Prochazkova & Kret, 2017), thus rendering observers unable to control their own pupil size (e.g., Laeng & Sulutvedt, 2014)—which in turn makes this an especially honest and reliable signal of one's attentional state. As a result, this cue has been used in hundreds of recent studies (of everything from memory and decision making to language and emotion), in experiments employing *pupillometry* (for reviews, see Binda & Murray, 2015; Laeng et al., 2012; Sirois & Brisson, 2014).

An extensive body of research has thus employed pupil size as a *dependent measure*. Here, in contrast, we employ it as a *stimulus* for social perception. If the apprehension of pupil size is so helpful to scientists, might it be similarly helpful to us in everyday life? Might the human visual system be naturally and automatically engaging in 'unconscious pupillometry'? Past work has shown that when viewing faces, pupil dilation influences neural processing even when observers do not notice such differences: faces with dilated pupils, for example, elicit greater amygdala activity (Amemiya & Ohtomo, 2012; Demos et al., 2008), though this phenomenon is not always observed (see Harrison et al., 2006). But might the detection of others' pupil size also influence awareness and behavior, even when we are not conscious of faces (much less eyes or pupils) in the first place? Here, in what is to our knowledge the first investigation of the perception of pupil size in social vision, we asked whether dilated pupils are automatically prioritized in visual processing, even outside of conscious awareness.

5.3 EXPERIMENT 1: PUPILS VS. BUTTONS

Observers viewed displays featuring faces whose pupils were artificially dilated or constricted, as depicted in Fig. 5.1a. We used continuous flash suppression (CFS; Tsuchiya & Koch, 2005;

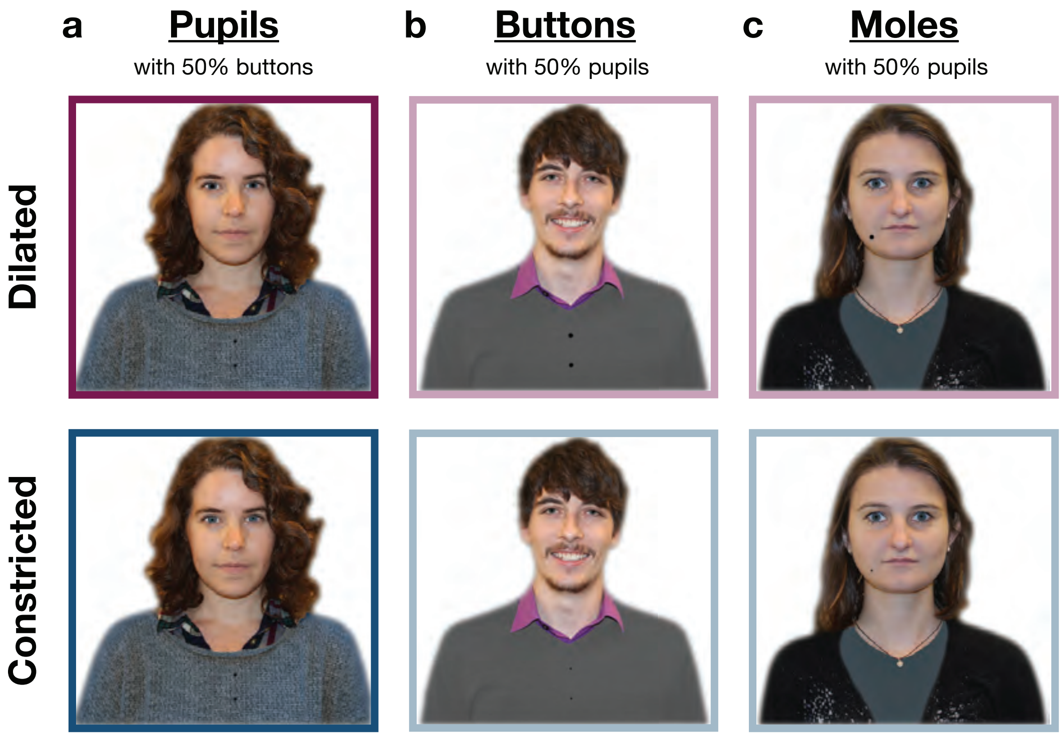

Figure 5.1: Stimuli from Experiments 1, 2, and 3. (a) Sample face with Constricted and Dilated pupils used in Experiments 1 and 2. (b) Sample face with Constricted and Dilated buttons used in Experiment 1. (c) Sample face with Constricted and Dilated moles used in Experiment 3.

for a review, see Stein, 2019) to render these faces invisible, and measured the time they took to break through interocular suppression (Fig. 5.2a). To ensure that any difference between dilated and constricted pupils was due to the perception of others' attention, per se (rather than to lower-level physical differences between the stimuli), we employed a control condition in which the same physical manipulation (i.e., large vs. small black dots) was applied to identical stimuli that lacked social significance: buttons on the actors' shirts (as depicted in Fig. 5.1b). And additional features of the experimental design further ruled out explanations that appeal to differences in contrast or spatial attentional biases.

5.3.1 METHOD

OBSERVERS

Thirty members of the Yale/New Haven community (21 females; average age = 21.93 years, *SD* = 3.99 years) participated in exchange for course credit or monetary compensation. Since to our knowledge no previous studies have employed pupil size in measures of visual awareness, this sample size was determined arbitrarily—but this was done before data collection began, was preregistered, and was fixed to be identical in both of the pupil-dilation experiments reported here. An additional two observers whose average accuracy was below 80 % were removed from further analyses and replaced, per the preregistered exclusion criteria. All experimental methods and procedures were approved by the Yale University Institutional Review Board.

APPARATUS

Stimuli were presented on a Dell 2208WFPT monitor with a 60 Hz refresh rate, using custom software written in Python with the PsychoPy libraries (Peirce et al., 2019). Observers placed their head in a chinrest and viewed the display through a custom-made mirror haploscope. The display was 90cm away, and subtended approximately 29.51° × 18.68° (with all extents reported below based on this distance).

STIMULI

As in the example depicted in Fig. 5.1, photographs of four individuals (2 males and 2 females) were taken and each was further modified according to the following procedure: the background was removed, the silhouette was feathered, the iris was lightened and/or pigmented (e.g., blue or green), and the shirt color was modified to roughly match the iris color. The pupils were then manipulated to be small (approximately 0.05° × 0.05°), medium (approximately 0.09° × 0.09°), or large (approximately 0.13° × 0.13°). The ensemble of the two pupils was then rotated 90° and pasted on the shirt to create buttons of the same size. As a result, several different versions of these images were created featuring (1) medium pupils with medium buttons, (2) small pupils with medium buttons, (3) large pupils with medium buttons, (4) medium pupils with small buttons, and (5) medium pupils with large buttons (see Fig. 5.1a and Fig. 5.1b). (For additional sample stimuli, see the online Supplementary Data.)

The functional part of the display consisted of two vertically centered 11.99° × 15.87° regions centered 7.53° to the left and right of the screen center. Each had a gray (#6E6E6E) background and a centered fixation dot (radius = 0.32°) with a black (#000000) inside and a red (#C72819) outline (stroke width = 0.14°), and was surrounded by a frame filled with static noise to support binocular alignment (0.81° stroke), and an outer red border (#C72819; 0.09° stroke).

Eighty Mondrian masks were created, each consisting of 1500 circles positioned randomly within the left-hand region (as depicted in Fig. 5.2a), each with a different radius (randomly selected from 0.18° to 1.26°) and color (randomly selected between white [#FFFFFF], yellow [#FFFF00], fuchsia [#FF00FF], red [#FF0000], lime [#00FF00], aqua [#00FFFF], blue [#0000FF], and black [#000000]).

PROCEDURE

At the beginning of each trial, observers saw the frames and fixation dots, and (if necessary) adjusted the haploscope mirrors until the left and right regions were binocularly fused. They then pressed a key to start the trial, after which the Mondrian masks immediately began flashing at 10 Hz on a randomly selected side. The face (7.18° × 6.87°, 7.18° × 7.01°, 7.18° × 7.15°, or 7.18° × 7.22°) was shown on the other side (horizontally centered within the frame and vertically displaced 4.22° either above or below the center of the frame), with its opacity linearly increased from 0 % to its maximum opacity over the course of the first second. As soon as observers saw any part of the image emerge into their awareness, they immediately indicated its position with respect to the fixation dot by pressing either the up or down arrow key. The trial ended after a response, or after 8 s had elapsed—at which point the next trial immediately began.

Design

Observers completed 2 blocks of 96 trials each (2 sizes [small/large] × 2 items [pupils/buttons] × 2 positions within the frames [up/down] × 4 identities × 3 repetitions), for a total of 192 trials. The trial order was randomized for each observer, and there were four self-paced breaks evenly spaced throughout the experiment. The experimental trials were preceded by 16 trials featuring different stimuli (license plates). The first 4 were practice trials, the results of which were not recorded. The remaining 12 functioned as a pre-test: observers were excluded from moving on to the experimental trials if their accuracy was below 75 % or if their average reaction time was below 1.0 s. Following these 16 practice trials, observers completed a staircasing procedure aimed at determining their optimal fade-in opacity. These trials featured the same faces as in the main experiment, but with medium-sized pupils and buttons. The opacity was initially 50 %, and was updated on each trial (in steps of 20 %, 10 %, 10 %, and 5 %) until observers responded accurately and within 3 s on 2 of the last 4 trials.

5.3.2 RESULTS AND DISCUSSION

Trials were removed from further analyses according to the following preregistered criteria: (1) missed (4.97/192 on average); (2) inaccurate (3.33/192 on average); and (3) more than 2 standard deviations away from each observer's mean (9.83/192 on average). The average breakthrough times for small and large items are depicted in Fig. 5.2b, separately for pupils and buttons. Inspection of this figure reveals that breakthrough times were faster for large vs. small pupils, but not for large vs. small buttons. Statistical analyses confirmed a reliable difference between large and small pupils (2.38 vs. 2.47 s, $t(29) = 2.50$, $p = .018$, $d_z = 0.46$), no difference between large and small buttons (2.50 vs. 2.48 s, $t(29) = 0.70$, $p = .490$, $d_z = 0.13$), and a reliable interaction ($t(29) = 2.45$, $p = .020$, $d_z = 0.45$). Thus, faces with dilated pupils enter awareness faster than faces with constricted pupils, and this difference vanishes when the same stimuli are presented in a socially meaningless way (as shirt buttons).

5.4 EXPERIMENT 2: PUPIL DILATION AND ATTRACTIVENESS?

The effects obtained in Experiment 1 demonstrate that faces with dilated pupils gain preferential access into visual awareness, even controlling for lower-level visual factors. Although we were motivated to test such effects by the well-established connection between pupil dilation and heightened attention, past work has also uncovered links between dilated pupils and other overt social impressions. Perhaps most notoriously, pupil dilation has been associated with perceived attractiveness: adult male observers have been reported to judge female faces to be more attractive when their pupils are dilated (Gründl et al., 2012; Hess, 1965, 1975; for a review, see Laeng & Alnæs, 2019). Might the prioritization for faces with dilated pupils observed in Experiment 1 thus be mediated by their increased attractiveness, rather than perceived attention per se? This possibility is supported, in principle, by prior results indicating

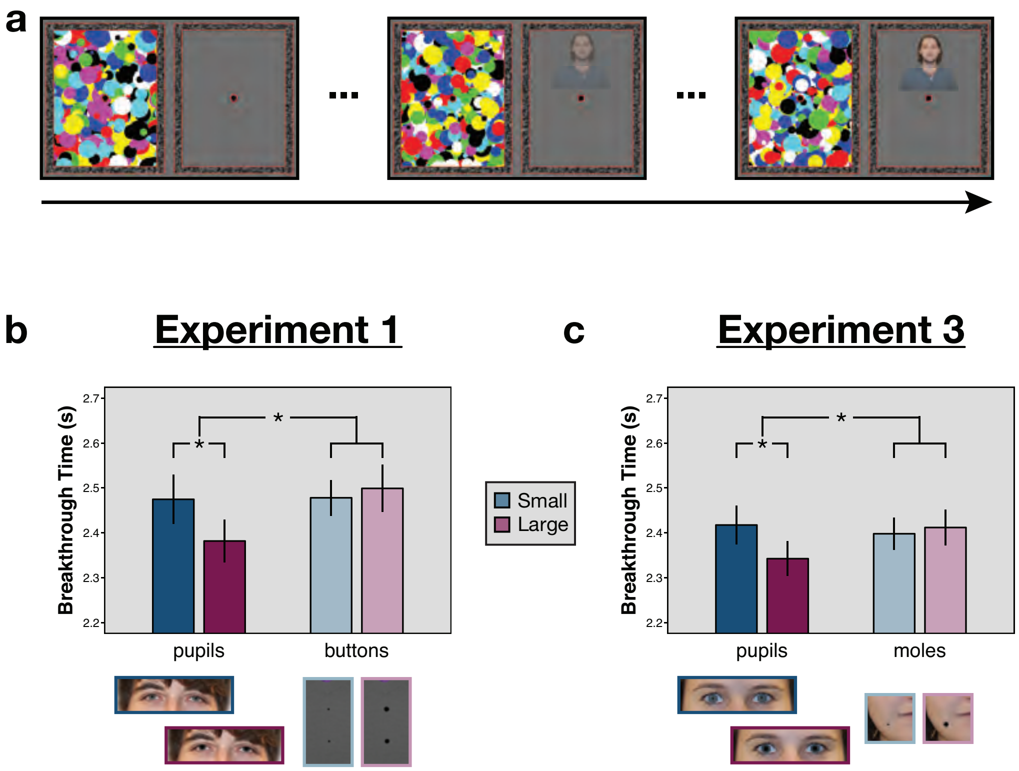

Figure 5.2: Methods and results from Experiments 1 and 3. (a) Depiction of the Continuous Flash Suppression paradigm (see text for details). (b) Stimuli close-up and average breakthrough times for dilated vs. constricted pupils and buttons (Experiment 1). (c) Stimuli close-up and average breakthrough times for dilated vs. constricted pupils and moles (Experiment 3). Error bars indicate 95 % confidence intervals, subtracting out the shared variance.

that more attractive faces gain preferential access into visual awareness (Hung et al., 2016; Jiang et al., 2006; Nakamura & Kawabata, 2018). To find out, we tested whether the same faces used in Experiment 1, presented under the same viewing conditions, would be perceived as differentially attractive depending on pupil dilation. If these stimuli were to be rated as more attractive with dilated (vs. constricted) pupils, then that would be consistent with the possibility that the results of Experiment 1 might have reflected an 'attractiveness effect' rather than the perception of heightened attention.

5.4.1 METHOD

OBSERVERS

1304 new observers (649 females; average age = 35.73 years, *SD* = 13.66 years) were recruited through Prolific Academic (www.prolific.co), and each completed a single trial in a 1-2min session in exchange for monetary compensation. (All observers resided in the U.S., had at least a 95 % Prolific approval rate, had previously completed at least 100 Prolific tasks, and had normal or corrected-to-normal acuity.) This sample size was determined and preregistered before data collection began based on an a-priori power analysis, which suggested that 1302 observers would suffice to achieve 95 % power to detect a conventionally small effect size (Cohen's $d = 0.20$) with a .05 alpha level. This number was then rounded up to 1304 in order to have an equal number of observers for each image.

APPARATUS

After agreeing to participate, observers were redirected to a website where stimulus presentation and data collection were controlled via custom software written in HTML, JavaScript, PHP, and CSS. (Since the experiment was rendered on observers' own web browsers, viewing distance, screen size, and display resolutions could vary dramatically, so we report stimulus dimensions below using pixel [px] values.)

STIMULI

Observers viewed the same stimuli employed in Experiment 1 in the Constricted and Dilated pupils conditions (small pupils with medium buttons, and large pupils with medium buttons; see Fig. 5.1a). The functional part of the display consisted of a 430 × 420 px region centered in their browser window and with a gray (#6E6E6E) background, thus matching the image background in Experiment 1.

PROCEDURE AND DESIGN

Each observer viewed a photograph of a single person (one of four possible identities; 230 × 220 px, 230 × 225 px, 230 × 229 px, or 230 × 232 px; roughly 7.2° × 6.9°, 7.2° × 7.0°, 7.2° × 7.2°, or 7.2° × 7.2°) centered in their browser window and within the functional part of the display. These sizes were chosen to match those of Experiment 1: (a) average viewport size was approximated using a sample of 400 participants from another online study conducted a few months prior (Colombatto et al., 2020, Experiment 3b; median viewport width = 1350 px); (b) average distance from the monitor was approximated using the average arm's length (~60 cm); and (c) average display size was approximated using a standard 20 inch diagonal and a 16:9 aspect ratio. Observers' browser windows were automatically put in full-screen mode at the beginning of the experiment, and observers were asked to sit at arm's length from the monitor. They were then instructed to view the image as carefully as possible, as it would be displayed only once. The photograph was displayed upon a keypress and after a 0.5 s delay. To match the viewing conditions from Experiment 1, the image was presented with its opacity linearly increased from 0 % to its maximum opacity over the course of the first second (with the maximum opacity set at 53 %, which was the average fade-in opacity from observers in Experiment 1 as determined by the staircasing procedure they completed prior to beginning the experiment). The image was then displayed at full opacity for an additional 1.4 s (such that the total presentation time was 2.44 s, matching the average response time in Experiment 1). After a 0.5 s delay, observers were then asked to rate how attractive that person looked. To respond, they simply clicked on one of nine buttons, numbered 1 through 9, with 1 labeled as "definitely not attractive" and 9 labeled as "definitely attractive". They then answered questions that allowed us to exclude observers (according to the preregistered criteria) who encountered technical problems (n = 14; e.g., reporting that "my trackpad accidentally got clicked and it went to the next page" during the instructions), or who misremembered the instructions as indicated on a multiple-choice question (n = 39; e.g., mis-reporting that they were supposed to rate the photograph on its perceived trustworthiness or competence, rather than its attractiveness). We also removed observers whose browser windows were smaller than 500 × 500 px (n = 4). The resulting unique excluded observers (n = 55, some of whom triggered multiple criteria) were replaced without ever analyzing their data. This design resulted in a total of 8 images (2 pupil sizes [small/large] × 4 identities), and each was viewed by 163 unique observers.

5.4.2 RESULTS

Pupil dilation did influence attractiveness judgments, but not in the predicted direction: faces with large pupils were reliably judged as less attractive than those same faces with small pupils (4.81 vs. 5.03, $t(1302) = 2.37$, $p = .018$, $d = 0.13$). Because prior reports of increased perceived attractiveness for pupil dilation were mostly based on male observers only (e.g., Hess, 1965), we also conducted an additional exploratory analysis testing whether the effect of pupil dilation on attractiveness judgments might be modulated by observers' gender. For the purposes of this analysis, we only analyzed data from observers who identified as "Female" (n = 649) or "Male" (n = 632), excluding those who selected "Other" (n = 20) or "I'd rather not say" (n

= 3). A two-way between-subjects ANOVA on the attractiveness ratings from the remaining 1281 observers revealed a main effect of pupil size ($F(1, 1277) = 4.51$, $p = .034$, $\eta_g^2 = 0.004$), no main effect of observer gender ($F(1, 1277) = 0.39$, $p = .533$, $\eta_g^2 < 0.001$), and no interaction ($F(1, 1277) = 0.03$, $p = .860$, $\eta_g^2 < 0.001$).

5.4.3 Discussion

This experiment was designed to investigate an alternative explanation for the results of Experiment 1—namely that prioritization into visual awareness might be driven by higher perceived attractiveness of faces with dilated pupils—rather than an effect of heightened perceived attention. On the contrary, however, the results revealed a small yet reliable effect wherein faces with dilated pupils in this stimulus set were rated as *less* attractive, and an additional analysis confirmed that this effect did not interact with participant gender. Far from providing an alternate explanation for the results observed in Experiment 1, the current results thus suggest that the previous experiment may even have been *underestimating* the effect of perceived attention—since the slight (if robust) attractiveness difference due to pupil dilation observed here was in the *opposite* direction from that consistent with the initial experiment's results (while previous CFS studies have consistently found that more attractive faces are prioritized for entry into visual awareness, as cited above). We did not find this reversed effect to be especially surprising, however. In fact, despite the long-held belief that faces with dilated pupils are perceived as more attractive by adult men (Hess, 1965), many studies have repeatedly failed to observe this purported effect (e.g., Amemiya & Ohtomo, 2012; Demos et al., 2008; Hicks et al., 1967) or have observed it only inconsistently (e.g., Tombs & Silverman, 2004).

5.5 Experiment 3: Pupils vs. Moles

The buttons employed as a control stimulus in Experiment 1 were identical to the pupils while lacking social significance—but of course they also differed in their location (i.e., appearing on the shirt instead of the face). In this experiment, we thus employed a control stimulus that also appeared on the face (often very near to the eyes or mouth) yet lacked social meaning: moles (as depicted in Fig. 5.1c).

5.5.1 Method

This experiment was identical to Experiment 1 except as noted here.

Thirty new observers (18 females; average age = 21.27 years, *SD* = 3.67 years) were recruited, with this preregistered sample size chosen to exactly match that of Experiment 1. An additional two observers whose average accuracy was below 80 % were removed from further analyses and replaced.

The stimuli employed in Experiments 1 and 2 were modified such that all buttons were removed, and one of the pupils was instead pasted onto the person's face (at a different location for each distinct person, always slightly above or below the eyes or mouth) to create (what appeared to be) a high-contrast mole of the same size. (These 'moles' were placed near to the eyes and mouth since those are the regions that observers tend to fixate during free viewing, though recall that in this study the faces were rendered invisible.) As a result, several different versions of these images were created featuring (1) medium pupils with a medium mole, (2) small pupils with a medium mole, (3) large pupils with a medium mole, (4) medium pupils with a small mole, and (5) medium pupils with a large mole.

5.5.2 Results and Discussion

Trials were removed from further analyses according to the following preregistered criteria: (1) missed (7.80/192 on average); (2) inaccurate (3.37/192 on average); and (3) more than 2 standard deviations away from each observer's mean (10.20/192 on average). The average breakthrough times for small and large items are depicted in Fig. 5.2c, separately for pupils and moles. Inspection of this figure reveals that breakthrough times were faster for large vs. small pupils, but not for large vs. small moles. Statistical analyses confirmed a reliable difference between large and small pupils (2.34 vs. 2.42 s, $t(29) = 2.21$, $p = .035$, $d_z = 0.40$), no difference between large and small moles (2.41 vs. 2.40 s, $t(29) = 0.47$, $p = .645$, $d_z = 0.08$), and a reliable interaction ($t(29) = 2.58$, $p = .015$, $d_z = 0.47$).

These results thus fully replicated the advantage for dilated pupils found in Experiment 1, while also confirming that this difference does not depend on the specific contrast with shirt buttons. And this absence of unconscious prioritization for 'mole dilation' (i.e., when the dilated stimuli were deprived of social meaning) seems especially remarkable given that the 'dilated' moles were visually more salient than the 'constricted' moles (and to a degree that went beyond the contrast in the pupil stimuli); for example, they had an especially high contrast with the background skin, and they made the faces less symmetrical. And as discussed below in the General Discussion, other aspects of these experiments also rule out potential explanations based on differential contrast, or spatial attentional biases.

5.6 General Discussion

Despite its social significance, pupil dilation is an exceptionally visually subtle signal—since dilated vs. constricted pupils differ by just a fraction of a degree of visual angle. (You might notice that the two faces in Fig. 5.1a look rather remarkably—if somewhat ineffably—different, despite differing by only a few pixels.) Indeed, this difference was so subtle that the observers in our experiments almost never even overtly noticed the variations in pupil size, despite each seeing 192 images in Experiments 1 and 3. (In post-experiment debriefing, only 2 of the 60 observers [1 in each experiment] reported any awareness of this manipulation—one referring to a difference in eye color, and the other mentioning that "Some people['s eyes] looked more intense".) This degree of subtlety makes the key results of this study all the more striking: these few pixels of difference—only when seen as dilated pupils—automatically facilitated the entry of faces into visual awareness.

This effect cannot be explained by appeal to a greater degree of visual contrast between the irises and the pupils, for two reasons. First, the shirts in Experiment 1 were modified to roughly match the color of the iris. Second, the contrast between the clear skin and the moles in Experiment 3 was actually considerably higher than that between the pigmented irises and the pupils—such that a contrast-based explanation would have to predict a greater dilated-vs.-constricted effect for moles compared to pupils.

This effect also cannot be explained by appeal to biases of spatial attention (for example if observers are generally biased to attend to the positions in which eyes appear), for three reasons. First, the positions of the faces were randomized so that on each trial they could appear either on the top or bottom of the display—such that there was no single region where the eyes appeared. Second, the positions in which the eyes appeared actually differed dramatically (by up to 28 px, or 0.38° × 0.32°) across the four separate identities featured in the experiments. And third, an explanation which appealed to spatial biases (relating to where the eyes were expected to appear) would predict that the effects should be greater when the same identity (and thus the identical eye positions) happened to repeat from one trial to the next—but if anything, the opposite was the case. (Collapsing across Experiment 1 and 3, the constricted-vs.-dilated effect was unreliable for repeated identities [$t(59) = 1.07$, $p = .290$, d_z = 0.14] but was robust for different identities [$t(59) = 2.51$, $p = .015$, $d_z = 0.32$], with no reliable interaction [$t(59) = 0.44$, $p = .658$, $d_z = 0.06$].)

The results of Experiment 2 further suggest that the influence of pupil dilation on visual awareness is not mediated through differences in perceived attractiveness. Rather, we suggest that these results reflect a more direct form of unconscious "attention to attention"—such that faces who seem to have heightened attention are prioritized in visual awareness. This is consistent with the fact that pupil dilation has been associated not only with attractiveness (inconsistently!), but with many other forms of heightened attention and arousal—including those due to emotionally arousing pictures (Bradley et al., 2008) and to difficult decisions (Lavín et al., 2014). In this way, the current results complement other recent findings that faces looking directly at (or turned towards) the observer break into awareness faster than do faces looking (or turned) away (Chen & Yeh, 2012; Gobbini et al., 2013; Stein et al., 2011). Whereas those previous studies can be interpreted as demonstrations that human visual processing is especially sensitive to *whether* others are attending to us (as signaled by directed gaze), the results of the current study indicate that visual processing is also sensitive to far subtler degrees of perceived attention (as signaled by pupil size), even with direct gaze.

In short, the current results suggest that the perceived attentional state of others can in turn cause us to attend to them—a novel form of 'attentional contagion'.

6

Conclusion

The work presented in this dissertation was motivated by the simple observation that others' eyes are the most important stimuli in our environment, but I conclude by suggesting that they might not be so special after all: Perception instead seems to be driven by deeper properties such as others' perceptual and attentional states, which are extracted and integrated into the very way we visually experience the world. This theoretical view is empirically demonstrated via two complementary approaches: In studies of 'minds without eyes', we demonstrate that the effects of eye gaze such as 'distracting stares' and 'stares in the crowd' can arise from eye-less stimuli, when they nonetheless signal others' attention (Chapter 2 and Chapter 3); and conversely, in studies of 'eyes without minds', we demonstrate that the effects of eye gaze are reduced when the eyes do not signal any underlying pattens of attention and intentions, even though they clearly look like eyes, as in the phenomena of 'gaze deflection' and 'unconscious pupillometry' (Chapter 4 and Chapter 5). In these concluding notes, I discuss our work more broadly (including several other empirical projects we have conducted that complement these chapters in a larger research program), as well as its rich connections to other areas of psychology.

6.1 Minds without Eyes

Chapter 2 investigates a notorious effect of eye contact, wherein memory for simple geometric shapes is disrupted by the presence of someone looking at you (Nie et al., 2018; Wang & Apperly, 2017). But this 'distracting' effect of eye gaze obtained even with simple 'mouth' shapes that were seen to be directed at (vs. away from) observers, despite not resembling eyes at all, and even after controlling for lower-level properties. Chapter 3 demonstrates that the well-known 'stare-in-the-crowd' effect also obtains with cone stimuli that are not just eyeless, but in fact look nothing like faces at all—and yet are nonetheless also readily perceived as being directed toward or away from the observer. These investigations of 'minds without eyes' demonstrate that effects of eye gaze might not be so eye-specific after all, since they obtain with stimuli that are entirely eyeless (as in the 'mouths' of Chapter 2) and even faceless (as in the 'cones' of Chapter 3).

Critically, the effects obtained in these experiments suggest that the apprehension of the direction of others' attention and intentions might be truly extracted in perceptual processing, for several reasons. First, percepts of intentionality in the current studies seem to depend on subtle details of the displays, e.g., the precise orientation of mouths in Chapter 2, and the coordinated orientation of cones in Chapter 3. Second, they seem automatic, insofar as participants were engaged in primary tasks (e.g., change detection, visual search), and it seems unlikely that any of our participants decided to prioritize directed mouths or cones: If anything, ignoring the mouths would have helped detect changes in the shapes in Chapter 2. As such, these experiments exhibit classic signatures of perceptual processing (Beck, 2018; Scholl & Gao, 2013), and suggest that putative effects of 'eye contact' might be better characterized as effects of 'mind contact', wherein the apprehension of attention and intentions seems to truly be extracted in perception.

6.1.1 Hidden Intentions

In additional work, we have also applied a similar methodological approach to what is perhaps the most remarkable influence of others' gaze: that on unconscious processing. Past work employing continuous flash suppression (CFS; see Chapter 5 for a description) has demonstrated that faces looking at you break into conscious awareness faster than faces looking away (Chen & Yeh, 2012; Stein et al., 2011). In a paper with Ben van Buren and Brian Scholl, we demonstrate that this prioritization also obtains with mouth stimuli (as in those employed in Chapter 2), and with cone stimuli (as in those employed in Chapter 3; Colombatto, van Buren, & Scholl, under review; see Fig. 6.1 for sample stimuli and results). In additional experiments, we rule out possible confounding factors such as differential symmetry or response biases (Moors et al., 2019). These effects of 'hidden intentions' not only demonstrate that another class of eye contact effects might in fact have little to do with eyes per se, but are especially striking in that they suggest that the perception of attention and intentions can operate in the absence of visual awareness.

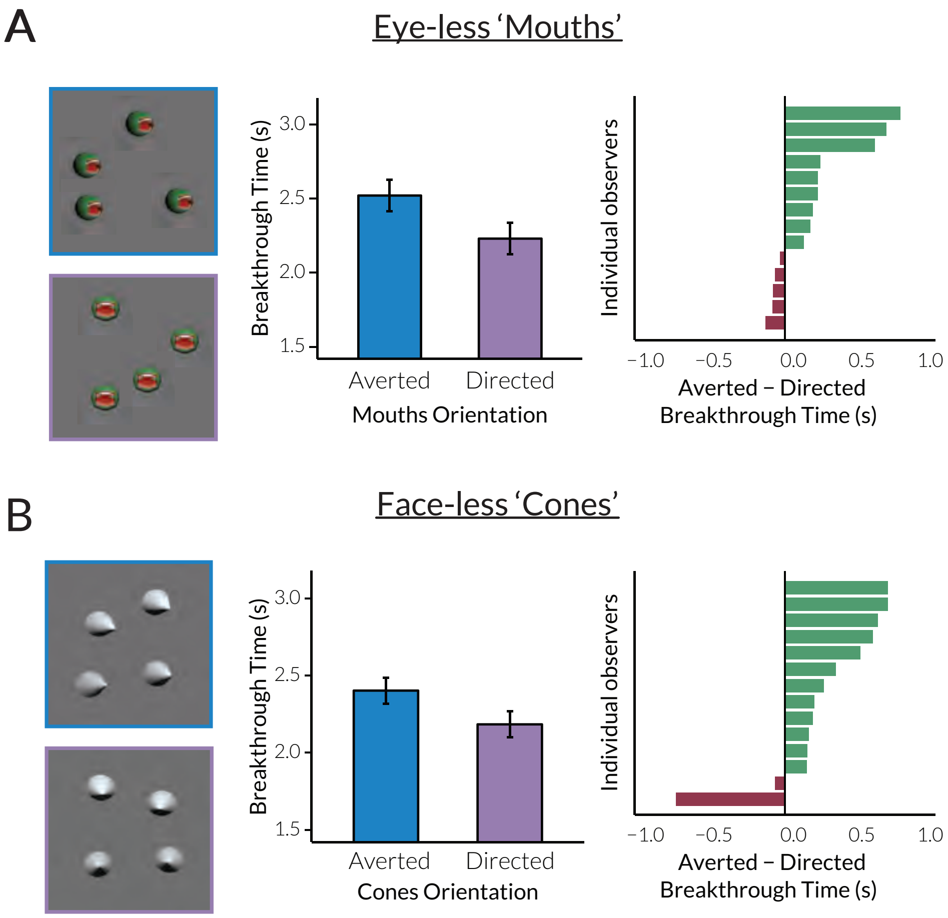

Figure 6.1: Sample stimuli and results from experiments of 'hidden intentions'. (a) Sample Averted and Directed displays employed in experiments of 'hidden intentions' using mouth stimuli, along with results averaged across subjects, and results depicted separately for each subject. (b) Sample Averted and Directed displays in experiments of 'hidden intentions' using cone stimuli, along with results averaged across subjects, and results depicted separately for each subject. Error bars reflect 95 % confidence intervals (CIs), subtracting out the shared variance.

6.1.2 Agency without Agents?

The logic of the 'minds without eyes' approach critically requires that stimuli such as mouths or cones are indeed perceived as minds, meaning that their direct and averted orientations are perceived as intentional. This may seem like an odd suggestion, since after all an inanimate object may happen to be directed towards us just by chance (e.g., when we face the shore at a particular angle and a wave seems to come towards us). In the current studies however, directedness was conveyed not through the orientation of one stimulus, but through the coordinated orientation of *many* stimuli. And unlike the orientation of a single stimulus, the coordinated orientation of multiple stimuli is unlikely to result from chance: Our minds instead automatically infer an agentic force at play. Of course, consistency does not always imply agency: If we see a bouquet of sunflowers in a vase and they are all facing towards the sun, we don't infer that they are agentic, since their coordinated orientation can be solely explained by natural forces. But if all sunflowers all point towards *us*, especially if they are placed in different locations, then natural forces can hardly explain the consistency—and we rather automatically infer the presence of other agents, in a sort of 'agency without agents'.

This intuition is especially apparent in the 'land art' movement, where artists create compositions by rearranging materials typically found in the environment such as rocks, leaves, or branches. Born as an ecological movement against the commercialization of art-making, land art aims exactly at demonstrating how humans can intervene in the landscape and leave their trace using solely natural materials (see Fig. 6.2 for some examples, and especially Fig. 6.2b, which is reminiscent of the 'cone' stimuli from Chapter 3). And indeed, this idea is not just a cultural movement, but traces back to the very origins of our civilization, as revealed by prehistoric monuments (e.g., Stonehenge) which document the presence of agents while being distinctly non-agent-like. In fact, we might even have evolved to detect traces of other agents via visual inspection of the environment, while other animals can rely on their heightened sense of smell. But beyond its cultural and historic significance, the phenomenon of 'agency without agency' is incredibly common in our everyday lives. Imagine returning home after it snowed and seeing footsteps leading up to your doorway: Someone already got home! Or

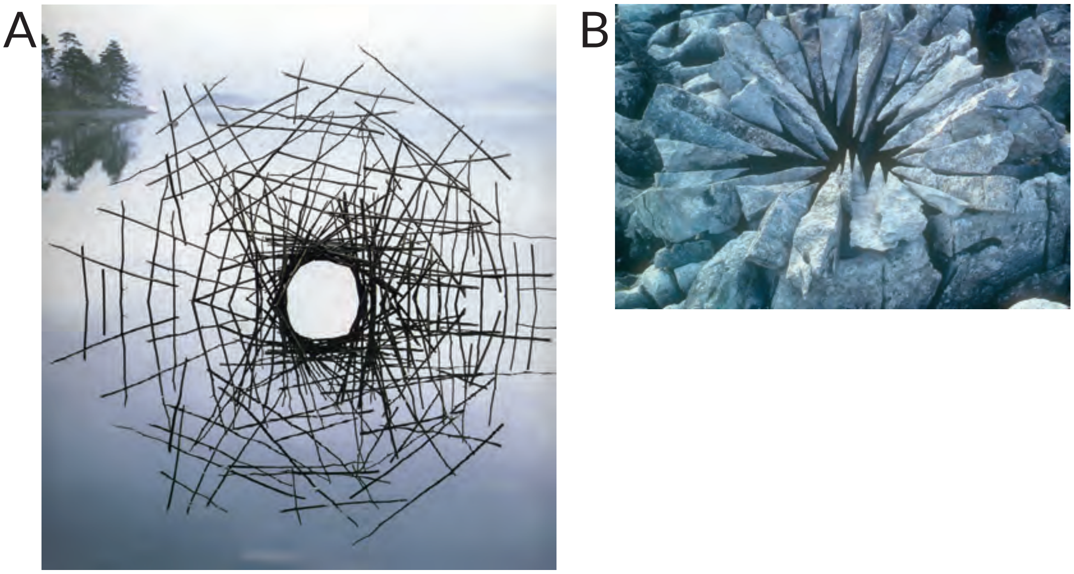

Figure 6.2: Sample land art. (a) "Screen" by Andy Goldsworthy (1998). (b) "Hole Covered with small pointed Rocks" by Andy Goldsworthy, Clapham, Yorkshire (UK, 1980).

imagine hiking along a trail and noticing a stack of rocks along the side of the path: Someone was there before you! So while social perception typically employs stimuli that are explicitly agent-like (e.g., faces, bodies, eyes), it seems clear that agency can be perceived even in stimuli that look nothing like agents, e.g., in configurations that are unlikely to have resulted from chance or natural forces.

6.2 Eyes without Minds

While Chapter 2 and Chapter 3 demonstrate that effects of eye gaze can arise from cues to others' intentions even when they are eye-less, Chapter 4 and Chapter 5 adopt a complementary approach: Might the effects of eye gaze be reduced when the eyes do not signal any underlying patterns of intention? Chapter 4 employs the common (yet previously unexplored) phenomenon we called 'gaze deflection' to demonstrate that gaze cueing is reduced for gazes that are not directed at anything in particular—even when controlling for lower-level spatial and temporal confounds. This result seems particularly surprising because gaze cueing had previously been thought to be not just dependent on the eyes per se (as many other effects of gaze), but especially automatic and reflexive, such that "any stimulus possessing eye-like attributes can trigger spatial orienting of attention" (Itier & Batty, 2009, p. 847). In Chapter 5, we further demonstrate this sensitivity to minds (vs. eyes) in what is perhaps an even more direct way: Faces rendered invisible through CFS enter awareness faster when they seem to be attentive such that their pupils are dilated. These effects of 'unconscious pupillometry' demonstrate that the perception of eye contact is sensitive to whether others are indeed paying attention, and also provide further evidence that cues to others' attention can be processed even in the absence of visual awareness (as in the 'hidden intentions' described in Section 6.1.1). Overall, these investigations of 'eyes without minds' thus demonstrate that 'mind contact' might really drive the effects of eye gaze, since it is the perception of attentional states rather than the physical features per se that can modulate effects as powerful as gaze cueing (as in Chapter 4), and as foundational as unconscious eye contact (Chapter 5).

And just as in 'minds without eyes', these investigations of 'eyes without minds' truly seem to implicate perceptual processes: It seems highly unlikely that any of our participants decided to not follow deflected gazes, or to prioritize dilated pupils. In fact, not only did we ensure that participants were unaware of the purpose of our experiments, but they also seemed unaware of the manipulations themselves. For example, only 2 out of the 60 participants in Chapter 5 guessed that the eyes might have been changing in the images, and none pointed to the pupils specifically. And indeed, in CFS paradigms participants are even unaware that images are presented in the first place! The effects reported here also seem to be independent from any *judgment* of gazing direction: For example, in gaze deflection observers are of course able to tell which direction the person who got caught turns towards, and yet attention responds otherwise. Similarly, all faces in studies of unconscious pupillometry were facing forward, and yet unconscious processing was modulated by subtle if not ineffable cues to their attention. In this way, the perception of attention seems to depend on subtle details of the displays—and not just in terms of how they look, but also in terms of how they unfold over time: It would look extremely odd if not ridiculous for a person who got caught staring to turn away extremely slowly! And even beyond their speed, movements in gaze deflection unfold in a very particular sequence: (1) Person A is staring at person B, then (2) B looks at A, and (3) A quickly turns away. Impressions of gaze deflection in this sense seem to be highly dependent on stimulus properties—and indeed the effects are weaker in control experiments where this sequence is altered (e.g., when A turns away before B looks at her).

6.2.1 Social Reordering

Of course, stimulus dependence by itself does not necessarily implicate perception: If anything, it seems natural that we first perceive these brute properties (e.g., temporal order), and later construct impressions of gaze deflection as higher-level explanations (e.g., for why events might have unfolded in that way). But in recent work in collaboration with Yi-Chia Chen and Brian Scholl, we have demonstrated that this is not always true: Social impressions are not just constructed from lower-level properties, but they are extracted early enough that they can alter the extraction of those (seemingly more primary) properties in the first place. This work was inspired by findings in the physical domain wherein impressions of causality can alter impressions of temporal order; for example, when viewing animations wherein a block (A) moves towards a second block (B), and when they are adjacent A stops and B starts moving, participants tend to report an impression that block A caused block B to move. But such 'causal' impressions also lead observers to misperceive block A as having moved *before* block B, even when in fact they start moving at the same time (Bechlivanidis & Lagnado, 2013, 2016; Tecwyn et al., 2020). In a social version of this 'reordering' effect, we showed participants animations where person A is staring at person B, and then A and B turn simultaneously (as depicted in Fig. 6.3a). Across seven experiments, we found that participants consistently misperceived the temporal order of events in such animations with A as turning before B, consistent with gaze deflection (as depicted in Fig. 6.3b). This illusion held across different paradigms such as spontaneous descriptions and forced selections, and even after controlling for lower-level properties of the displays. And it was also especially powerful, since it occurred not just when A and B were turning at the same time, but also when A actually turned *after* B, demonstrating that gaze deflection does not just alter, but can even *reverse* the perceived order of events.

This illusion demonstrates that others' attentional states can be extracted in perception itself, for several reasons. First, the illusion vanishes when the animations are replaced with descriptions of the events, suggesting that the effect critically depends on the visual input, rather than inferences about what might be most likely to have happened (a 'blindfold test'; van Buren & Scholl, 2018, under review). Second, it seems especially automatic, since participants are simply told to watch the animations, with no mention of temporal order at all. Third, it seems to result from a distinct phenomenology, as is clear from the descriptions participants generate (e.g., reporting that "the woman on the left then turns her head to look to the left", as opposed to a typical participant in the control condition "both women glanced to their right simultaneously"). But most importantly, this illusion demonstrates that social impressions are not just hierarchically constructed from simple visual features, but can also (perhaps counterintuitively) alter what those properties look like in the first place.

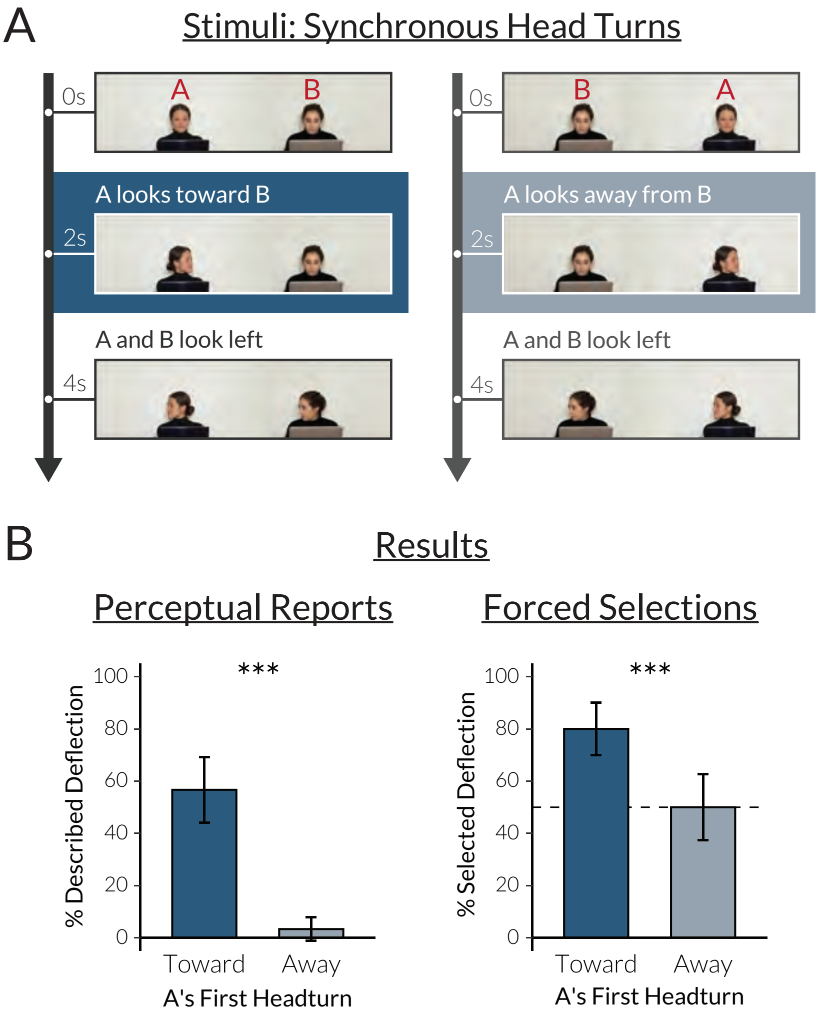

Figure 6.3: Sample stimuli and results from experiments of 'social reordering'. (a) A schematic depiction of the animations observers viewed, including Toward-B animations and Away-from-B animations. (b) Percentage of observers who perceived B as turning earlier than A for Toward-B vs. Away-from-B animations as revealed by their descriptions (left panel) or choices (right panel). Error bars indicate 95 % CIs.

6.2.2 Gazing to Look vs. Gazing to Think

In addition to its implications for the roots of various eye gaze effects, and for the perceptual nature of the extraction of others' attention, this work introduces a new type of question in the study of eye gaze—focused not on *where* someone is looking, but rather on *whether* people are attentive in the first place. In short, some gazes are simply just not worth following—either because people are not intending to look in those locations (as in Chapter 4), or because they are distracted (as in Chapter 5).

In some more recent work, we pushed this question even further: Sometimes people might be highly engaged, and yet their gaze is still not worth following because they are not attending to the environment—but rather to their own thoughts and memories. Attention can in fact be directed at information coming in from our environment (e.g., when listening to another person), but also at information that is already in our minds (e.g., when later remembering that conversation; Chun, Golomb, & Turk-Browne, 2011). But despite being so fundamentally different, external and internal attention may manifest in similar behaviors. For example, if someone asks us to count the lights on the ceiling, we will shift our gaze to look at the ceiling—while of course also attending there. But if they ask us how many lights were on the ceiling of our childhood bedroom, we may also shift our gaze toward the ceiling—not because we are attending there, but because it helps us focus on our memory. This is a well-known tendency known as 'gaze aversion' (Doherty-Sneddon et al., 2002; Glenberg, Schroeder, & Robertson, 1998): During cognitively demanding tasks such as arithmetic reasoning or memory retrieval, both children and adults tend to look in locations that are less visually cluttered (e.g., the ceiling, the floor, a blank wall). This indeed results in better performance, presumably because disengagement from the external world aids internal focus (Doherty-Sneddon & Phelps, 2005). Beyond being a fascinating phenomenon in itself, gaze aversion is a rare case where someone may shift their gaze in a way that is not quite worth following, since of course they are not gazing *to look*, but rather gazing *to think*.

Might our visual system then be sensitive not just to whether someone is attentive (vs. distracted), but also to whether *the target* of their attention is external vs. internal? In a collaboration with Yi-Chia Chen and Brian Scholl, we showed participants videos of an actress who is initially looking forward; she is then asked a question, and before responding she gazes upward and to the side. As she is gazing away, a symbol quickly flashes in the direction of her gaze—and we later tested participants' discrimination accuracy for this symbol (see Fig. 6.4a for a schematic of the animation). The critical manipulation was the nature of the question the actress was asked, which motivated her to attend either externally ("Who painted that piece of art on the wall over there?"), or internally ("Who painted that piece of art we saw in the museum?"). In this way, all the animations were visually identical, and only the audio track changed. As depicted in the results in Fig. 6.4b, across multiple samples and multiple pairs of 'external' vs. 'internal' questions, gaze cueing was far weaker when the gaze was directed internally. These results suggests that vision is not just sensitive to whether someone is attentive (vs. distracted, as in Chapter 5), or to whether they are attending to the location they are gazing at (vs. another location, as in Chapter 4). Rather, it is also sensitive to whether others are in fact attending to their environment, such that we follow people when they gaze to look, but not when they gaze to think.

6.3 Related Perspectives

These investigations collectively help address a long-standing question in cognitive and social psychology about the nature of social vision—namely about whether it is driven by the saliency of the visual stimuli per se, or whether it might be driven instead by richer social impressions. In particular, the major contribution of the 'mind contact' proposal is the suggestion that not only are the effects of gaze driven by the apprehension of others' mental states, but that these are automatically extracted in visual processing itself. And while this possibility has been rarely explored in vision science, it has rich connections to related work

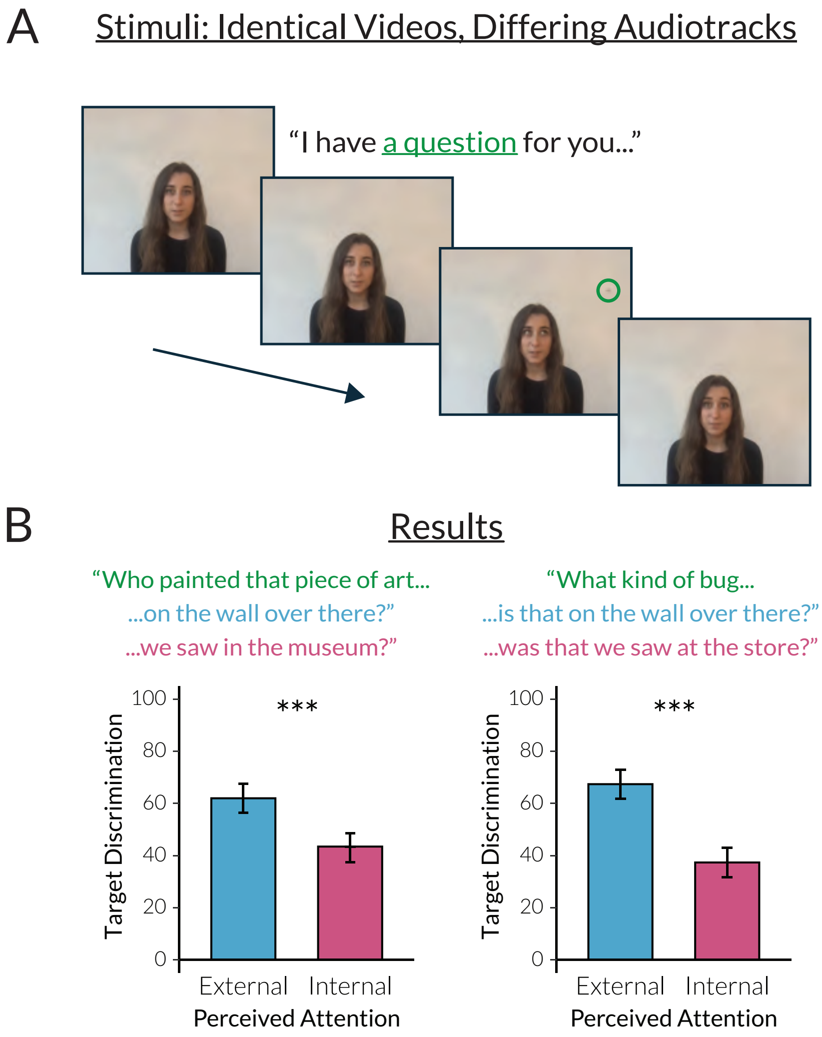

Figure 6.4: Sample stimuli and results from experiments of 'gazing to think vs. gazing to look'. (a) A schematic depiction of the animations observers viewed. (b) Average accuracy in the letter discrimination task for External vs. Internal questions. Error bars indicate 95 % CIs.

in other areas of our field that have instead posited a mentalistic understanding of gaze. In the remainder of this discussion, I will briefly review related work in vision science suggesting that perception constructs more abstract representations of gaze from the eyes, and then turn to related ideas in phylogeny, ontogeny, and neuroscience, which have implicated more mentalistic inferences in the perception of gaze.

6.3.1 Psychophysics: Coding Gaze from Eyes

While vision science rarely explores the sorts of higher-level percepts of mental states investigated in this dissertation, recent work suggests that perception might construct representations of gaze that incorporate more than just merely the direction in which eyes are pointing. For example, some work has shown that the perception of gazing direction integrates multiple cues in addition to the eyes themselves, including head and body orientation (Otsuka, Mareschal, & Clifford, 2016; for a review, see Langton, Watt, & Bruce, 2000). And when observers are repeatedly exposed to faces wherein gazing direction is incongruent with head orientation (in the so-called 'Wollaston' faces; Wollaston, 1824), perceptual adaptation follows the perceived rather than the veridical gazing direction (Palmer & Clifford, 2018; for a review, see Clifford & Palmer, 2018). The perception of averted gaze also seems to be influenced by the presence of objects in the surrounding context, such that estimates of gazing direction are biased towards nearby objects (Lobmaier, Fischer, & Schwaninger, 2006).

Similarly, gaze is not just perceived as direct when it is truly direct (i.e., when it falls directly on the vector from the gazers' eyes to the observers' eyes), but is rather gradually perceived as less direct the more it deviates from 'true' direct—a range known as the 'cone of gaze' (Gamer & Hecht, 2007). The cone of gaze is highly sensitive to a variety of factors including perceptual uncertainty (Clifford et al., 2015; Mareschal et al., 2013) and viewing distance (Gamer & Hecht, 2007; Vine, 1971), but also stimulus properties such as expressions of happiness (Lobmaier & Perrett, 2011) or anger (Ewbank, Jennings, & Calder, 2009; Harbort et al., 2013), and observer characteristics such as social anxiety (Gamer et al., 2011; Jun et al., 2013; Schulze, Renneberg, & Lobmaier, 2013) and stress (Rimmele & Lobmaier, 2012).

Collectively, this work suggests that "representations of other people's gaze direction in the visual system [...] are abstracted from lower-level facial cues", and rely instead on high-level perceptual representations (Palmer & Clifford, 2018, p. 82). While of course none of these effects implicate the apprehension of mental states, they are indeed deeply connected to the 'mind contact' perspective in that perception itself might construct more abstract representations that just computing eye direction. And while vision science has considered perception to at most integrate different directional cues, work in other areas of psychology has explored the possibility that our responses to others' gaze might integrate an understanding of their minds.

6.3.2 Phylogeny: A Reinterpretation of Gaze

One of the most remarkable facts about gaze is that an incredible range of species beyond humans exhibit particular responses to direct (vs. averted) gaze, e.g., by escaping or freezing upon eye contact. These include fish (e.g., jewelfish: Coss, 1979), reptiles (e.g., iguanas: Burger, Gochfeld, & Murray Jr, 1992; snakes: Burghardt & Greene, 1988; or lizards: Hennig, 1977), birds (e.g., chickens: Gallup, Cummings, & Nash, 1972), and rodents (e.g., mice: Topal & Csanyi, 1994; for a review, see Emery, 2000). In fact, several species even follow others' gaze—including dogs (Hare, 2002), goats (Kaminski et al., 2005), dolphins (Tschudin et al., 2001), ravens (Bugnyar, Stöwe, & Heinrich, 2004), seals (Scheumann & Call, 2004), and penguins (Nawroth, Trincas, & Favaro, 2017). And gaze following has also been demonstrated in most primates, including lemurs (Ruiz et al., 2009; Shepherd & Platt, 2007), New World monkeys (e.g., marmosets: Burkart & Heschl, 2007), and Old World monkeys (e.g., rhesus macaques: Emery et al., 1997; Tomasello et al., 1998). Gaze following in monkeys is also modulated by social status, such that dominant monkeys will follow the gaze of other dominant monkeys, but not subordinates (Shepherd, Deaner, & Platt, 2006), just like dominance of the face cue modulates gaze cueing in humans (Jones et al., 2010).

While responses to eye gaze are pervasive across several animal species, closer investigations have highlighted an evolutionary shift that sets primates (and humans in particular) apart. Despite resulting in similar behaviors, gaze perception in primates seems to rely on qualitatively different cognitive mechanisms (Rosati & Hare, 2009), wherein gaze is represented not just as a behavior, but as a reflection of the minds behind the eyes. This 'reinterpretation' hypothesis (Povinelli & Vonk, 2004) is supported by two main lines of evidence. First, humans adjust their behavior in the presence of eyes on the basis of reputational concerns, e.g., by cooperating more in economic games (Bateson et al., 2006; Dear, Dutton, & Fox, 2019; cf. Cai et al., 2015; Northover et al., 2017)—but no such effects have been found in other primates (Nettle, Cronin, & Bateson, 2013). This divergence is commonly taken to suggest that humans, but not other primates, understand that the eyes signal someone is present and is watching (Grossmann, 2017). (Although I note that of course the lack of behavioral effects from watching eyes in primates could simply reflect a lack of reputational concerns.)

A perhaps more convincing line of evidence is that primates seem not just to be sensitive to where others are looking, but also to have a sense that gaze reflects an underlying state of 'seeing'. For example, chimpanzees and monkeys will look for an alternative target of gaze if an experimenter looks at an object they have already seen with a surprised expression (Drayton & Santos, 2017; MacLean & Hare, 2012), and they will reach for food that a competitor can't see (Flombaum & Santos, 2005; Hare et al., 2000). Chimpanzees (but not monkeys) will also move around barriers to see what another person is gazing towards (Itakura, 1996; Povinelli & Eddy, 1996, 1997; Tomasello, Hare, & Agnetta, 1999), and use others' gazing patterns to learn about contingencies of food rewards (Anderson, Montant, & Schmitt, 1996; Itakura & Tanaka, 1998).

Of course, these sophisticated behaviors might simply reflect learnt contingencies, without necessarily involving any underlying state of seeing: "The learned social cues remain subordinate to cues that tend to covary with someone 'seeing' them, but have no bearing on 'seeing' [... such that] behavioral abstractions formed by chimpanzees are essentially postural heuristics that have nothing to do with 'seeing' at all" (Povinelli & Barth, 2005, p. 217). On the other hand, others have attributed these sophisticated behaviors to the fact that primates can *think* about others' minds (e.g., Flombaum & Santos, 2005), and "reason about the knowledge and perceptions of others" (Santos, Nissen, & Ferrugia, 2006, p. 1176). Of course, reasoning about mental states is fundamentally different than the type of visual experiences implicated in the 'mind contact' perspective. But nonetheless, the upshot from this vast comparative literature is that gaze perception has evolved from a mere reflexive response to the eyes, to a more sophisticated phenomenon that seems sensitive to the underlying states of seeing. In short, a shift from eyes to minds.

6.3.3 Ontogeny: Starting with Eyes and Ending with Minds

This evolutionary shift from eyes to minds is paralleled (on a much accelerated timeframe) in human ontogeny. Gaze perception in humans is extremely sophisticated from the very beginning of life: 48 hours after birth, infants already preferentially orient to faces with direct gaze (Farroni et al., 2002), prefer to gaze at a face with the eyes visible (Batki et al., 2000), and show a rudimentary form of gaze cueing (Farroni et al., 2004; although see below for a critical role of motion). At around seven weeks of age, infants start fixating mostly on the eye region when scanning faces (Haith et al., 1977), and by the age of three months they direct their attention towards the direction of adults' gaze (Hood, Willen, & Driver, 1998). The powerful effects of eye contact on other cognitive processes are also already present in infancy: Eye contact enhances face memory in 4-month-old infants (Farroni et al., 2007), and this effect increases over development from 6 to 11 years (Smith et al., 2006).

Infants' preferences for eyes and eye contact are of course compatible with a view of gaze perception as merely driven by the saliency of eyes as visual stimuli: Some for example claim that "gaze-cued attention is initially based on gross luminance information about eye direction", and "gaze-following should therefore be impossible if the relevant luminance information is removed" (Doherty, McIntyre, & Langton, 2015, p. 73). But infants will soon also develop more a sophisticated understanding of gaze (del Bianco et al., 2019). By 12 months of age, they will be surprised if adults turn to look towards empty locations (Csibra & Volein, 2008), and they will move around a barrier to see what others are looking at, instead of simply following the geometric line of gaze and looking at the barrier itself (Moll & Tomasello, 2004). By 14 months of age, they cease to overtly follow the gaze of unreliable lookers (Chow, Poulin-Dubois, & Lewis, 2008); by 18 months of age, they cease to follow head turns of adults with closed eyes (Brooks & Meltzoff, 2002, 2005), or those wearing opaque blindfolds (Meltzoff & Brooks, 2008; Senju et al., 2011). And remarkably, they also understand that gazing often signifies wanting: By 6 months of age, infants will be surprised if an agent reaches to grab a toy different from the one she had just gazed towards (Woodward, 1998). After viewing videos of a puppet attempting to climb a hill (the 'climber'), infants prefer to interact with a puppet who helped (vs. hindered) the climber (Hamlin et al., 2007), but only when the climber was gazing towards the hill—suggesting they understand that gazing direction corresponds to a goal to climb the hill (Hamlin, 2015). And conversely, social relevance can increase the saliency of gaze shifts from inanimate objects: 18-month-olds even follow the gaze of a metal robot, but only after it has engaged with adults in social interchanges such as reciprocal imitation (Meltzoff et al., 2010). Similarly, infants will follow the orientation of a non-face like robot (Movellan & Watson, 1987), a novel object (Johnson et al., 1998), and non-human objects (Deligianni et al., 2011) after watching them produce reactions contingent to their own movements or those of other objects. And indeed, the integration of eyes and intentions seems to underlie many forms of learning in infants, e.g., when they use gaze to disambiguate between possible objects of utterances (Baldwin, 1991), in a sort of 'natural pedagogy' (Csibra & Gergely, 2009).

The contribution of mental state understanding for the development of gaze perception is also supported by evidence from Autism Spectrum Disorder (ASD), a developmental disorder characterized by severe impairments in social communication and interactions (American Psychiatric Association, 2013). Among a variety of other symptoms, ASD is marked by a consistent failure to orient to the eyes (Dalton et al., 2005; Pelphrey et al., 2002), and to initiate joint attention (Charman et al., 1997; Charman, 2000). Individuals diagnosed with ASD are able to discriminate gazing direction (Leekam et al., 1997) and will learn to follow gaze cues (Kylliäinen & Hietanen, 2004) but only when predictive (Ristic et al., 2005), and do not show any advantage for gaze cues preceded by direct gaze (Senju et al., 2003)—suggesting a role for mental state understanding in the development of reflexive gaze cueing. Recent work has also shown that eye looking in infants later diagnosed with ASD is actually initially intact, and rapidly declines between 2 and 6 months of age (Jones & Klin, 2013)—just when other infants begin to perceive the 'mind' behind the eyes. And conversely, administration of oxytocin generally promotes prosocial behavior in humans, but also results in increased fixations to the eye region (Auyeung et al., 2015; Guastella, Mitchell, & Dadds, 2008; for a review, see Heinrichs, von Dawans, & Domes, 2009), even in ASD patients (Andari et al., 2010).

Of course, just as for primates, the influence of mental state understanding on responses to gaze might not necessarily reflect sophisticated reasoning about minds, but simply learnt associations between agents and their gazing behavior (Moore & Corkum, 1994): Over the course of the first year of life, infants learn that adults' gaze is often associated with important objects or salient events, in a form of reinforcement learning (Deák et al., 2014; Ishikawa, Senju, & Itakura, 2020; Jasso et al., 2012; Triesch et al., 2006). In this way, gaze following results purely from automatic and implicit associations between gaze cues and interesting events, where the perception of others' minds seems superfluous. As Leslie and Frith put it, "the concept of seeing/not seeing [...] could be a purely geometric-causal notion involving the construction of an imaginary line in space between the eyes and their target. [...] A geometric-causal notion does not require knowledge of the *experience* of seeing" (Leslie & Frith, 1988, pp. 317–318). And of course, both mentalistic and behavioristic accounts of the ontogeny of gaze perception critically differ from the 'mind contact' perspective in that they do not implicate perception per se in the extraction of mental states. But overall, whether via learnt associations or via mentalistic reasoning, it seems clear that infants come to have rich and sophisticated responses to gaze: an ontogenetic shift from eyes to minds.

6.3.4 Neuroscience: Integrating Eyes with Minds

Another reason why our visual system is typically thought to be tuned to perceive the eyes (rather than the minds behind them) is putative specialized processing for eye direction in the superior temporal sulcus (STS). This selectivity was initially discovered in monkeys with single-cell recordings (Perrett et al., 1992), and later also in humans with neuroimaging experiments (Calder et al., 2007; Hoffman & Haxby, 2000; Puce et al., 1998; Wicker et al., 1998), intracranial recordings (Allison et al., 1999; McCarthy et al., 1999; Puce, Allison, & McCarthy, 1999), and lesion case studies (Akiyama et al., 2006; Samson et al., 2004). Critically, findings that other attentional cues (e.g., arrows) do not engage this system (Hietanen et al., 2006; Vuilleumier, 2002) have led some to suggest that the STS is "engaged especially when the stimuli are represented as eyes" (Kingstone et al., 2004, p. 271), such that "the eyes constitute a special stimulus [...because] we may have evolved neural mechanisms devoted to gaze processing" (Langton et al., 2000, p. 50).

Several lines of evidence however suggest that neural responses to gaze cues might instead reflect responses to minds more generally. First, responses in the STS from its posterior to anterior subregions are increasingly invariant to gaze-irrelevant features such as head view (Carlin et al., 2011), and posterior STS seems sensitive to the meaning of gaze beyond its visual appearance. For example, pSTS activity is modulated by whether faces are looking towards (vs. away from) a target (Mosconi et al., 2005; Pelphrey et al., 2003), or the participants themselves (Conty et al., 2007; Pelphrey, Viola, & McCarthy, 2004; Watanabe et al., 2006). Selectivity is also challenged by the fact that eye contact activates a number of regions including not just STS but also fusiform gyrus (e.g., George et al., 2001), medial prefrontal cortex (e.g., Kampe, Frith, & Frith, 2003), orbitofrontal cortex (e.g., Wicker et al., 2003), amygdala (e.g., Adolphs et al., 2005; Kawashima et al., 1999; Spezio et al., 2007), and even ventral striatum (Pfeiffer et al., 2014; Schilbach et al., 2010; for reviews, see George & Conty, 2008; Itier & Batty, 2009; Nummenmaa & Calder, 2009; Senju & Johnson, 2009). And conversely, the human STS has been implicated in various other cognitive functions, from audiovisual integration (Beauchamp et al., 2004) to intentionality (e.g., Gao, Scholl, & McCarthy, 2012; Saxe et al., 2004) and belief inference (Saxe, 2006), so much to be called 'the chameleon of the human brain' (Hein & Knight, 2008).

Recent unifying frameworks suggest instead that the STS might generally respond to biological motion (Pitcher & Ungerleider, 2021). This function may be supported in part by its anatomical connections, as revealed by tractography studies identifying a white matter pathway from early visual cortex (V1) to STS via motion-selective temporal areas (V5/MT) both in macaques (Boussaoud, Ungerleider, & Desimone, 1990; Ungerleider & Desimone, 1986) and humans (Gschwind et al., 2012). In addition to its anatomy, the functional properties of the STS seem optimal for motion detection: For example, the STS is visual-field invariant, as opposed to the strong contralateral biases in other visual areas (Finzi et al., 2021; Pitcher et al., 2020). In fact, these anatomical and functional dissociations from the ventral and dorsal visual pathways have pointed to the existence of a third visual pathway on the lateral surface of the brain involving V1, V5/MT, and STS, and engaged in the dynamic aspects of social perception (Pitcher & Ungerleider, 2021). This hypothesis of specialized processing for agentic motion is also supported by evidence that gaze following in newborns requires motion of the eyes (Farroni et al., 2004), and does not occur with static face cues. This integration of gaze with other cues to agency and intentions (including possible motion) in the STS is consistent with the 'mind contact' perspective, wherein gaze serves as cue to others' intentions, such that they always seem to be "looking into the future" (Chen, Colombatto, & Scholl, 2018).

6.3.5 Top-Down Mental State Attributions

While the phylogenetic, ontogenetic, and neuroscientific perspectives outlined so far do not directly implicate perception as in the 'mind contact' perspective, they do suggest that the perception of others' eyes is integrated with some understanding of others' minds. And indeed, a tradition in social perception has claimed that gaze perception can be influenced by our beliefs and judgments about what others can see. In fact, some have even claimed that effects such as eye contact or gaze following are not just influenced, but entirely driven by top-down attributions of mental states: "Participants make an assumption of what the agent knows; it is knowledge that is the important component that drives the effect rather than vision" (Cole & Millett, 2019, p. 1975). As such, gaze effects stem not from our visual experience as in the 'mind contact' perspective, but rather from "what the observer believes about the other's mental state" (Teufel, Fletcher, & Davis, 2010, p. 376): Eye contact arises from "the observer's knowledge of being the target of another individual's attention" (Myllyneva

& Hietanen, 2015, p. 107), and gaze cueing is a case where "mind perception constitutes a source of top-down modulation on attentional orienting" (Wiese et al., 2018, p. 852).

This perspective is in part supported by evidence that gaze cueing effects vanish when the face cues have their eyes closed (vs. open; Nuku & Bekkering, 2008, Expt. 1), or are obstructed by a dark rectangle (vs. sunglasses; see Fig. 6.5a; Nuku & Bekkering, 2008, Expt. 2). Similarly, gaze cueing is reduced when participants are informed that the face cues are wearing googles that participants know are opaque (vs. transparent; Morgan, Freeth, & Smith, 2018; Teufel, et al., 2010), or when a physical barrier is inserted between the face cue and the target, with depth cues indicating the face is behind the barrier (vs. in front; see Fig. 6.5c; Kawai, 2011). And in fact, gaze cueing is reduced when viewing a human-like mannequin (vs. a real human; Wiese et al., 2012), or a mindless robot (vs. a robot whose eyes are controlled by a human; Wiese et al., 2012). But while these results are typically taken to demonstrate that "beliefs about another person's ability to see [...] have strong top-down effects on gaze processing" (Teufel et al., 2009, p. 1276), directly "facilitating and shaping the way in which social signals are processed on a lower level" (Teufel, Fletcher, et al., 2010, p. 1277), others have suggested that they might rather only be indirect (e.g., affecting attention via heightened imagery; Langton, 2009). And indeed, these effects are typically found at longer latencies or in less reflexive paradigms, such that they might affect free viewing but not reflexive gaze following (Kingstone et al., 2019; Kuhn et al., 2018).

Most importantly, this framework seems incompatible with the most automatic effects of gaze on perception and attention such as those explored in this dissertation: Eye contact could not break CFS faster (Stein et al., 2011) as a result of deliberative processes, since those cannot operate in the absence of visual awareness; and similarly, reasoning would certainly prevent us from following the gaze of anti-predictive cues (e.g., Driver et al., 1999). In fact, some of the effects of eye gaze seem downright counter to our beliefs: We certainly don't

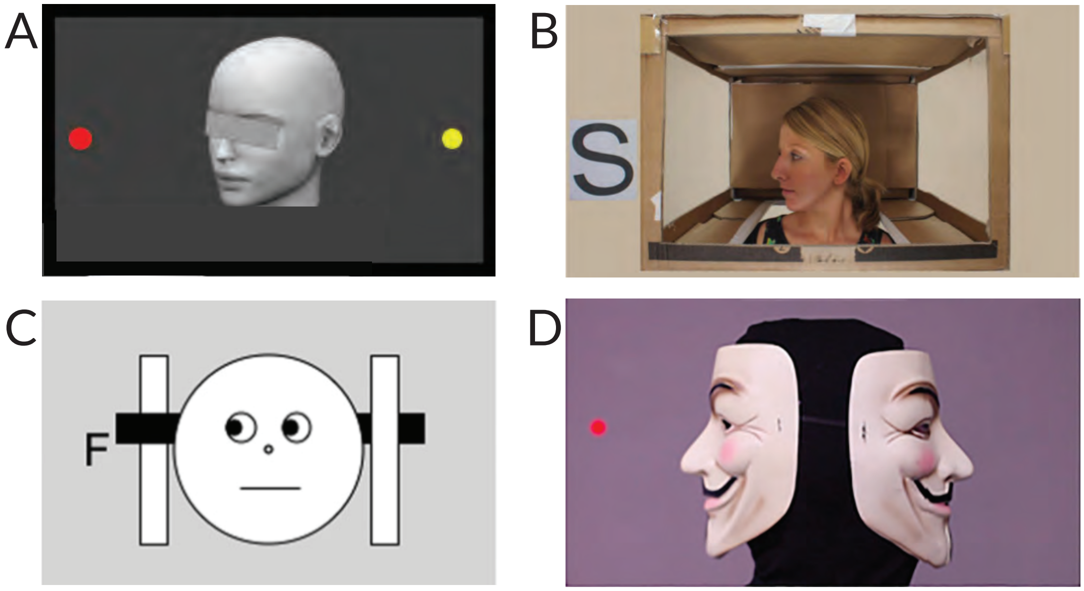

Figure 6.5: Sample stimuli from gaze cueing experiments manipulating whether the face cue can see the target. (a) Faces with their eyes blocked by a dark rectangle are compared to faces with their eyes covered by see-through sunglasses (Nuku & Bekkering, 2008, Expt. 2). (b) Faces in a box with its sides cut out are compared to faces in a box with its sides obstructing its view (Cole et al., 2015). (C) Faces in front of a barrier (as signaled by the black bar being in front of the barrier yet behind the face) are compared to faces behind a barrier (as signaled here by the black bar being behind both the barrier and the face; Kawai, 2011). (D) Masks covering the back of the head are compared to masks covering the face (Kingstone et al., 2019).

believe clouds and peppers can see, and yet we follow the gaze of pareidolia faces (Takahashi & Watanabe, 2013). So while this work has focused on our *beliefs* about others' attention, which fail to explain its most automatic and powerful effects, this discrepancy is resolved in the 'mind contact' perspective wherein others' attention does not just result from higher-level judgments of mental states, but is extracted in perception itself.

6.4 Apparent Challenges

The 'mind contact' perspective is based on the idea that gaze perception is driven by the apprehension of others' attention; but this seems incompatible with two lines of recent work—one demonstrating that gaze cueing effects can arise from face cues even when they are not looking at the target, and one that has found effects of 'minds' in cues that do not seem animate in any way.

6.4.1 Knowing vs. Seeing

While the 'mind contact' perspective is grounded on the idea that gaze effects are driven by others' attention, recent work has instead suggested that effects such as gaze cueing arise even when others are clearly not looking at anything. For example, when viewing an actor wearing two masks—one in the front and one in the back of their head, gaze cueing is the same regardless of whether the mask directed towards the target is the one covering the front vs. the back of the actor's head (Fig. 6.5d; Kingstone et al., 2019). How can gaze following be driven by the perception of others' attention, when we follow the back of people's heads (which if anything, is more often directed in the direction opposite of attention)? Similarly, in some cases gaze cueing has been found to be unaltered by the presence of physical barriers inserted between the face and the target (Fig. 6.5b; Cole et al., 2015). If the 'mind contact' perspective posits that another person's intentions drive the effects of gaze, then those should be weaker in these cases where the face cues are obviously not looking at the target. And these paradigms seem straightforward in this sense: Of course inserting a physical barrier between an agent and an object will result in the agent not being able to see the object, and of course an agent will not be able to see anything at all from the back of their head. And in fact, participants in these experiments are quite accurate at reporting whether the face cue can or cannot see the target.

But crucially, *knowing* that one cannot see does not necessarily imply that our phenomenology reflects such conditions, for both practical and theoretical reasons. First, some of these manipulations require quite some scrutiny and working memory (e.g., remembering whether you are viewing the back vs. the front of the head). Some manipulations are in fact implemented in a way that is downright counter to their indented function (e.g., the target in Fig. 6.5b actually appears to be obstructed by the corner of the box, despite this being the 'open box' condition, and does not fall along the line of gaze). But even beyond these practical considerations, it is also clear more generally that these paradigms investigate extremely uncommon situations: Rarely do we see agents stuck in boxes, or wearing masks on both sides of their heads. Even the manipulations that might seem more natural such as physical barriers appear strange: Rarely do we see agents suddenly turn to look at a blank barrier.

The unfamiliar nature of these cues is not problematic in itself, but rather because we tend to explain otherwise strange behaviors by ascribing intentions—the very intentions these manipulations were designed to eliminate. If we see someone staring at a blank barrier, of course we will think they cannot see objects behind it; but they might be staring exactly because they are interested in the objects behind it! This is indeed a common phenomenon: When you are sitting at the theater right before a show, you might stare at the closed curtain not because your perceptual system is broken, but because you are attending to the actors behind it in anticipation. Similarly, if you lost your keys you might gaze around the room and lay your eyes on various bags, purses, coats, jackets, pants, etc. that you might have left them in: And while of course you can't see through these items, you would certainly be attending to their content to figure out if the keys might be there. In fact, our tendency to ascribe intentionality to otherwise inexplicable behavior might even lead us to *infer* the presence of a target when there is none. For example, if we saw someone suddenly turn towards a new location, we might assume they heard a noise coming from there (even if they turned with their eyes closed). And similarly, if we saw someone stare at a wall from a different perspective than us, we might assume there is something interesting on the wall (e.g., a painting).

These considerations highlight how beliefs about what others can see only approximate our phenomenology of their attention and intentions—and it is thus unsurprising that manipulating *beliefs* about seeing does not influence how we visually experience agents in the first place. And indeed, implicating perception requires a completely different class of stimuli—one that is grounded in biological significance rather than in explicit beliefs participants might have. A wonderful example of this is the fact that pointing gestures trigger attentional shifts, but only when the cue is the index finger—and not the index finger and the middle finger together, the little finger, the little finger lengthened to match the index finger, or even the index finger shortened to match the little finger (Ariga & Watanabe, 2009). This 'index-finger advantage' is remarkable especially because participants are certainly able to judge the direction in which all these cues are pointing. This suggests that "visual processing for socially relevant and communicative stimuli can be so narrow-tuned that subtle differences in visual stimuli [...] have influences on reflexive attentional shifts" (Ariga & Watanabe, 2009, p. 264). And critically, this perceptual sensitivity is rooted in the social relevance of these stimuli rather than their lower-level properties, just like in the 'mind contact' view, the special status of others' eyes is rooted in their social relevance as cues to others' minds rather than in their visual properties per se.

6.4.2 Effects of Inanimate Objects

The possibility that perception itself might extract social properties is also challenged by evidence that stimuli that are typically taken to be non-social produce similar effects as gaze.

For example, cueing effects arise from gaze cues but also from arrows (Eimer, 1997; Kuhn & Kingstone, 2009; Ristic, Friesen, & Kingstone, 2002; Tipples, 2002), suggesting that the social significance of eyes might play no role in these effects. This work has also documented some differences between gaze and arrows cues, namely that gaze cues seem more irresistible (Friesen et al., 2004), rely on different neural mechanisms (Kingstone et al., 2004), and are selectively impaired by STS lesions (Akiyama et al., 2006) and in autism (Senju et al., 2004; for reviews, see Birmingham & Kingstone, 2009; Itier & Batty, 2009). But these differences have again been attributed to post-perceptual processing of biologically significant cues, rather than perception itself being well-tuned to social information.

The 'perceiving perception' framework we introduce has two implications for this theory of gaze cueing as a non-social process. First, the notion of 'minds without eyes' (see Section 6.1) suggests that arrows might be social, in the sense that their orientation is often not a coincidence. If we are biking down a street and see arrows on our lane, we would have no doubt they have been intentionally placed by an agent to indicate the direction of travel. And similarly, if a participant is completing an experiment and sees an arrow on the computer screen pointing at exactly 180°, they would have no doubt it had been intentionally programmed to point to the left, since that precise orientation is unlikely to occur by chance (for a similar argument on the social significance of arrows, see Kingstone et al., 2003). As such, it might be entirely plausible that inanimate cones or arrows would produce cueing effects when they appear to be non-random. Second, if the importance of gaze lies in its biological relevance as an honest cue to attention, then the eyes of a schematic faces are actually more akin to an arrow than to the eyes of real people—and so it is perhaps unsurprising that arrows and schematic faces give rise to the same effects. But this does not mean they would if compared to more biologically significant stimuli—and indeed, when looking at real world scenes containing both faces and arrows, people tend to focus on the people (Birmingham et al., 2008). In sum, it seems that apparent challenges to the 'mind contact' view simply arise from a misguided operationalization of others' minds, which influence our conscious experience not by virtue of our beliefs that someone is attending, and not by virtue of their superficial properties, but rather by virtue of subtle if not ineffable cues to others' minds that substantially alter our phenomenology.

6.5 A Creative Genius

The field of social perception encompasses a wide variety of topics, from the stereotypes we might have about other people (e.g., Freeman, Stolier, & Brooks, 2020), to how the presence of others might activate incidental knowledge or memory (e.g., Ferguson & Bargh, 2004). Ironically, most of this work has little to do with perception—which is instead implicated only insofar as the initial input to higher-level cognitive processes. This is perhaps because while it seems rather uncontroversial that we are able to draw inferences about others' mental states, the possibility that we might *perceive* attentional states is highly controversial, with some dismissing it as "definitely wrong" (McGinn, 1984, p. 123) if not "extremely paradoxical" (Price, 1931, p. 54). The claim that mental states are fundamentally imperceptible has longstanding roots in philosophy, starting with St. Augustine who claimed that "there is no way opened to our eyes to see the mind, a thing which cannot be seen by the eyes" (419, as cited in Miles, 1983). This assumption has historically led to the conclusion that knowledge of others' mental states must be indirect, and can only result from analogies to our own mental states (e.g., Russell, 1923) or as best explanation of observed behavior (e.g., Pargetter, 1984), or even just via testimony (Gomes, 2015).

While an epistemological discussion of the problem of other minds seems beyond the scope of this dissertation, it is worth noting that the imperceptible nature of other minds is endorsed to this date in psychology: “Mental states, and the minds that possess them, are necessarily unobservable constructs that must be inferred by observers rather than perceived directly” (Johnson, 2000, p. 22). Critically, this view is not just a theoretical position about *whether* other minds can be perceived in principle, but critically has far reaching implications for *how* mental states understanding might be achieved in practice, since it implies that perception plays no role in mental state understanding. In fact, leading theories of mental state understanding generally accept that mental states cannot be perceived, and are instead inferred through extra-perceptual steps, such as folk-psychological theories (Gopnik & Meltzoff, 1997; Premack & Woodruff, 1978), or simulations (Gordon, 1986; Heal, 1986).

But the experiments reported here paint a different picture of social perception, wherein perception is not merely an input to higher-level processes, but is rather well-tuned to extract social information such as the direction of others’ attention and intentions. This does not imply that intentions are physical entities contained in the input itself (a ‘strong’ view of direct perception; Lavelle, 2012). Additional inferences are required, but we argue that rather being carried out in additional post-perceptual processing, these inferences are extracted in perception itself, and integrated in the resulting conscious experience. While extremely uncommon in studies of social psychology, this view seems almost required from the perspective of vision science: The fact that retinal images are vastly underdetermined (e.g., they lack a 3D structure) and yet our percepts seem so rich (e.g., they do have a 3D structure) requires that additional processing is carried out prior to the construction of conscious experience. In this way, vision is not just a means for sensory transduction, but is rather a sophisticated process that transforms scarce input into rich conscious experiences via a series of ‘unconscious inferences’—a view that has a rich history in the study of perception (Pylyshyn, 1999; Rock, 1983; for a review, see Cavanagh, 2011).

These sorts of perceptual heuristics have been demonstrated to operate in a wide variety of contexts, supporting our percepts of spatial relationships (Ullman, 1984) and objecthood (Gregory, 1980; Scholl, 2001), and perhaps even more fundamentally determining what we take to be a stimulus in the first place (vs. a defect of our visual system; New & Scholl, 2008). More recently, unconscious yet sophisticated perceptual inferences have been shown to operate in domains that had been previously thought to be solely the province of higher-level though. These include biological motion (Chatterjee, Freyd, & Shiffrar, 1996), causality and animacy (for reviews, see Scholl & Gao, 2013; Scholl & Tremoulet, 2000), and also non-social ones such as physics (Firestone & Scholl, 2017, under review), history (Chen & Scholl, 2016), and even aesthetics (Chen et al., 2018). Overall, this research program highlights that perception is not just a means through which we extract inputs for higher-level processes, but a 'creative genius' (Hoffman, 2000, p. 1) that constructs rich and sophisticated experiences of the world around us including other people's minds—such that we 'perceive perception', and 'attend to attention'.

References

Adams, R., & Kleck, R. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science, 14,* 644–647.

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433,* 68–72.

Akechi, H., Senju, A., Uibo, H., Kikuchi, Y., Hasegawa, T., & Hietanen, J. K. (2013). Attention to eye contact in the West and East: Autonomic responses and evaluative ratings. *PLoS ONE, 8,* e59312.

Akiyama, T., Kato, M., Muramatsu, T., Saito, F., Umeda, S., & Kashima, H. (2006). Gaze but not arrows: A dissociative impairment after right superior temporal gyrus damage. *Neuropsychologia, 44,* 1804–1810.

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9,* 415–430.

Amemiya, S., & Ohtomo, K. (2012). Effect of the observed pupil size on the amygdala of the beholders. *Social Cognitive and Affective Neuroscience, 7,* 332–341.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: Author.

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 Sciences, 107,* 4389–4394.

Anderson, J. R., Montant, M., & Schmitt, D. (1996). Rhesus monkeys fail to use gaze direction as an experimenter-given cue in an object-choice task. *Behavioural Processes, 37,* 47–55.

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze. *Perception, 31,* 657–674.

Ando, S. (2004). Perception of gaze direction based on luminance ratio. *Perception, 33,* 1173–

1184.

Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception of where a face or television mydoubleqportrait is looking. *The American Journal of Psychology, 82,* 474–489.

Applegate, C. D., Kapp, B. S., Underwood, M. D., & McNall, C. L. (1983). Autonomic and somatomotor effects of amygdala central N. stimulation in awake rabbits. *Physiology & Behavior, 31,* 353–360.

Ariga, A., & Watanabe, K. (2009). What is special about the index finger?: The index finger advantage in manipulating reflexive attentional shift. *Japanese Psychological Research, 51,* 258–265.

Auyeung, B., Lombardo, M. V., Heinrichs, M., Chakrabarti, B., Sule, A., Deakin, J. B., … Baron-Cohen, S. (2015). Oxytocin increases eye contact during a real-time, naturalistic social interaction in males with and without autism. *Translational Psychiatry, 5,* e507–e507.

Böckler, A., van der Wel, R. P. R. D., & Welsh, T. N. (2014). Catching eyes: Effects of social and nonsocial cues on attention capture. *Psychological Science, 25,* 720–727.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62,* 875–890.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20,* 723–742.

Barbato, M., Almulla, A. A., & Marotta, A. (2020). The effect of trust on gaze-mediated attentional orienting. *Frontiers in Psychology, 11,* 1554.

Baron-Cohen, S. (1994). The Mindreading System: New directions for research. *Current Psychology of Cognition, 13,* 724–750.

Baron-Cohen, S. (1997). Mindblindness: An essay on autism and theory of mind. Cambridge, MA, US: MIT Press.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry, 38,* 813–822.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The mydoubleqreading the mind in the eyes test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42,* 241–251.

Barrett, H. C., Todd, P. M., Miller, G. F., & Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior, 26,* 313–331.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters, 2,* 412–414.

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development, 23,* 223–229.

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments. *Psychological Science, 17,* 514–520.

Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review, 13,* 1061–1066.

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron, 41,* 809–823.

Bechlivanidis, C., & Lagnado, D. A. (2013). Does the mydoubleqwhy tell us the mydoubleqwhen? *Psychological Science, 24,* 1563–1572.

Bechlivanidis, C., & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order. *Cognition, 146,* 58–66.

Beck, J. (2018). Marking the perception–cognition boundary: The criterion of stimulus-dependence. *Australasian Journal of Philosophy, 96,* 319–334.

Binda, P., & Murray, S. O. (2015). Keeping a large-pupilled eye on high-level visual processing. *Trends in Cognitive Sciences, 19,* 1–3.

Bindemann, M., Burton, A., & Langton, S. (2008). How do eye-gaze and facial expression interact? *Visual Cognition, 16,*708–733.

Birmingham, E., & Kingstone, A. (2009). Human social attention. *Annals of the New York Academy of Sciences, 1156,* 118–140.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Gaze selection in complex social scenes. *Visual Cognition, 16,* 341–355.

Bobak, A. K., & Langton, S. R. H. (2015). Working memory load disrupts gaze-cued orienting of attention. *Frontiers in Psychology, 6,* 1258.

Boussaoud, D., Ungerleider, L. G., & Desimone, R. (1990). Pathways for motion analysis:

Cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *Journal of Comparative Neurology, 296,* 462–495.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45,* 602–607.

Bradshaw, J. (1967). Pupil size as a measure of arousal during information processing. *Nature, 216,* 515–516.

Brodsky, S. L. (2013). Testifying in court: Guidelines and maxims for the expert witness, Second Edition. Retrieved May 2, 2021, from: https://www.apa.org/pubs/books/4317294

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38,* 958–966.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8,* 535–543.

Buchanan, H., Markson, L., Bertrand, E., Greaves, S., Parmar, R., & Paterson, K. B. (2014). Effects of social gaze on visual-spatial imagination. *Frontiers in Psychology, 5,* 671.

Bugnyar, T., Stöwe, M., & Heinrich, B. (2004). Ravens, Corvus corax, follow gaze direction of humans around obstacles. *Proceedings of the Royal Society B: Biological Sciences, 271,* 1331–1336.

Burger, J., Gochfeld, M., & Murray Jr, B. G. (1992). Risk discrimination of eye contact and directness of approach in black iguanas (Ctenosaura similis). *Journal of Comparative Psychology, 106,* 97.

Burghardt, G. M., & Greene, H. W. (1988). Predator simulation and duration of death feigning in neonate hognose snakes. *Animal Behaviour, 36,* 1842–1844.

Burkart, J. M., & Heschl, A. (2007). Understanding visual access in common marmosets, Callithrix jacchus: Perspective taking or behaviour reading? *Animal Behaviour, 73,* 457–469.

Buswell, G. T. (1935). How people look at pictures: A study of the psychology and perception in art. Univ. Chicago Press.

Butler, S. C., Caron, A. J., & Brooks, R. (2000). Infant understanding of the referential nature of looking. *Journal of Cognition and Development, 1,* 359–377.

Cai, W., Huang, X., Wu, S., & Kou, Y. (2015). Dishonest behavior is not affected by an image of watching eyes. *Evolution and Human Behavior, 36,* 110–116.

Caldara, R., Schyns, P., Mayer, E., Smith, M. L., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience, 17,* 1652–1666.

Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., & Henson, R. N. A. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology, 17,* 20–25.

Calder, A. J., Young, A., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 527–551.

Cameron, E. L., Tai, J. C., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research, 42,* 949–967.

Capozzi, F., & Ristic, J. (2020). Attention AND mentalizing? Reframing a debate on social orienting of attention. *Visual Cognition, 28,* 97–105.

Carlin, J. D., Calder, A. J., Kriegeskorte, N., Nili, H., & Rowe, J. B. (2011). A head view-invariant representation of gaze direction in anterior superior temporal sulcus. *Current Biology, 21,* 1817–1821.

Cavanagh, P. (2011). Visual cognition. *Vision Research, 51,* 1538–1551.

Chapman, H. A., & Anderson, A. K. (2012). Understanding disgust. *Annals of the New York Academy of Sciences, 1251,* 62–76.

Charman, T. (2000). Theory of mind and the early diagnosis of autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience (pp. 422–441). New York, NY, US: Oxford University Press.

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology, 33,* , 781–789.

Chatterjee, S. H., Freyd, J. J., & Shiffrar, M. (1996). Configural processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 916–929.

Chen, Y. -C., & Scholl, B. J. (2016). The perception of history: Seeing causal history in static shapes induces illusory motion perception. *Psychological Science, 27,* 923–930.

Chen, Y. -C., & Yeh, S. L. (2012). Look into my eyes and I will see you: Unconscious processing of human gaze. *Consciousness and Cognition, 21,* 1703–1710.

Chen, Y. -C., Colombatto, C., & Scholl, B. J. (2018). Looking into the future: An inward bias in aesthetic experience driven only by gaze cues. *Cognition, 176,* 209–214.

Chow, V., Poulin-Dubois, D., & Lewis, J. (2008). To see or not to see: Infants prefer to follow the gaze of a reliable looker. *Developmental Science, 11,* 761–770.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62,* 73–101.

Clifford, C.W.G., Mareschal, I., Otsuka, Y., & Watson, T. L. (2015). A Bayesian approach to person perception. *Consciousness and Cognition, 36,* 406–413.

Clifford, Colin W. G., & Palmer, C. J. (2018). Adaptation to the direction of others' gaze: A review. *Frontiers in Psychology, 9,* 2165.

Cline, M. G. (1967). The perception of where a person is looking. *The American Journal of Psychology, 80,* 41–50.

Cole, G. G., & Millett, A. C. (2019). The closing of the theory of mind: A critique of perspective-taking. *Psychonomic Bulletin & Review, 26,* 1787–1802.

Cole, G. G., Smith, D. T., & Atkinson, M. A. (2015). Mental state attribution and the gaze cueing effect. *Attention, Perception, & Psychophysics, 77,* 1105–1115.

Colombatto, C., & Scholl, B. J. (in press). Unconscious pupillometry: An effect of 'attentional contagion' in the absence of visual awareness. *Journal of Experimental Psychology: General.*

Colombatto, C., Chen, Y.-C., & Scholl, B. J. (2020). Gaze deflection reveals how gaze cueing is tuned to extract the mind behind the eyes. *Proceedings of the National Academy of Sciences, 117,* 19825–19829.

Colombatto, C., van Buren, B., & Scholl, B. J. (2019). Intentionally distracting: Working memory is disrupted by the perception of other agents attending to you—even without eye-gaze cues. *Psychonomic Bulletin & Review, 26,* 951–957.

Colombatto, C., van Buren, B., & Scholl, B. J. (2020). Gazing without eyes: A 'stare-in-the-crowd' effect induced by simple geometric shapes. *Perception, 49,* 782–792.

Colombatto, C., van Buren, B., & Scholl, B. J. (under review). Hidden intentions: Visual awareness prioritizes perceived attention even without eyes or faces.

Conty, L., Gimmig, D., Belletier, C., George, N., & Huguet, P. (2010). The cost of being watched: Stroop interference increases under concomitant eye contact. *Cognition, 115,* 133–139.

Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia, 45,* 3024–3037.

Conty, L., Tijus, C., Hugueville, L., Coelho, E., & George, N. (2006). Searching for asymmetries in the detection of gaze contact versus averted gaze under different head views: A behavioural study. *Spatial Vision, 19,* 529–545.

Cooper, R. M., Law, A. S., & Langton, S. R. H. (2013). Looking back at the stare-in-the-crowd effect: Staring eyes do not capture attention in visual search. *Journal of Vision, 13,* 1–22.

Coss, R. G. (1979). Delayed plasticity of an instinct: Recognition and avoidance of 2 facing eyes by the jewel fish. *Developmental Psychobiology, 12,* 335–345.

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE, 8,* e57410.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107,* 705–717.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13,* 148–153.

Csibra, G., & Volein, Á. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology, 26,* 1–11.

Dalmaso, M., Castelli, L., & Galfano, G. (2020). Social modulators of gaze-mediated orienting of attention: A review. *Psychonomic Bulletin & Review, 27,* 833–855.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., ... Davidson, R. J. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience, 8,* 519–526.

Davies, T. N., & Hoffman, D. D. (2002). Attention to faces: A change-blindness study. *Perception, 31,* 1123–1146.

Deák, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Developmental Science, 17,* 270–281.

Deaner, R. O., Shepherd, S. V., & Platt, M. L. (2007). Familiarity accentuates gaze cuing in women but not men. *Biology Letters, 3,* 64–67.

Dear, K., Dutton, K., & Fox, E. (2019). Do 'watching eyes' influence antisocial behavior? A systematic review & meta-analysis. *Evolution and Human Behavior, 40,* 269–280.

del Bianco, T., Falck-Ytter, T., Thorup, E., & Gredebäck, G. (2019). The developmental origins of gaze-following in human infants. *Infancy, 24,* 433–454.

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-months-old infants. *Developmental Psychology, 47,* 1499–1503.

Demos, K. E., Kelley, W. M., Ryan, S. L., Davis, F. C., & Whalen, P. J. (2008). Human amygdala sensitivity to the pupil size of others. *Cerebral Cortex, 18,* 2729–2734.

Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory & Cognition, 33,* 727–733.

Doherty-Sneddon, G., Bonner, L., & Bruce, V. (2001). Cognitive demands of face monitoring: Evidence for visuospatial overload. *Memory & Cognition, 29,* 909–919.

Doherty-Sneddon, Gwyneth, Bruce, V., Bonner, L., Longbotham, S., & Doyle, C. (2002). Development of gaze aversion as disengagement from visual information. *Developmental Psychology, 38,* 438–455.

Doherty, M. J., & Anderson, J. R. (1999). A new look at gaze: Preschool children's understanding of eye-direction. *Cognitive Development, 14,* 549–571.

Doherty, M. J., McIntyre, A. H., & Langton, S. R. H. (2015). Developmentally distinct gaze processing systems: Luminance versus geometric cues. *Cognition, 137,* 72–80.

Doi, H., & Ueda, K. (2007). Searching for a perceived stare in the crowd. *Perception, 36,* 773–780.

Doi, H., Ueda, K., & Shinohara, K. (2009). Neural correlates of the stare-in-the-crowd effect. *Neuropsychologia, 47,* 1053–1060.

Dovidio, J. F., & Ellyson, S. L. (1982). Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. *Social Psychology Quarterly, 45,* 106–113.

Downing, P., Dodds, C., & Bray, D. (2004). Why does the gaze of others direct visual attention? *Visual Cognition, 11,* 71–79.

Drayton, L. A., & Santos, L. R. (2017). Do rhesus macaques, Macaca mulatta, understand what others know when gaze following? *Animal Behaviour, 134,* 193–199.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6,* 509–540.

Duchenne (de Boulogne), G.-B. (1990). The mechanism of human facial expression. Cambridge, UK: Cambridge University Press.

Dunham, P. J., Dunham, F., & Curwin, A. (1993). Joint-attentional states and lexical acquisition at 18 months. *Developmental Psychology, 29,* 827.

Eimer, M. (1997). Uninformative symbolic cues may bias visual-spatial attention: Behavioral and electrophysiological evidence. *Biological Psychology, 46,* 67–71.

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion, 11,* 860–865.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17,* 124–129.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior, 6,* 238–252.

Elekes, F., & Király, I. (2021). Attention in naïve psychology. *Cognition, 206,* 104480.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews, 24,* 581–604.

Emery, N., Lorincz, E., Perrett, D., Oram, M., & Baker, C. (1997). Gaze following and joint attention in rhesus monkeys (macaca mulatta). *Journal of Comparative Psychology, 111,* 286–293.

Ewbank, M. P., Jennings, C., & Calder, A. J. (2009). Why are you angry with me? Facial expressions of threat influence perception of gaze direction. *Journal of Vision, 9,* 16–16.

Farkas, L. G. (1994). Anthropometry of the head and face. Lippincott Williams & Wilkins.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, 99,* 9602–9605.

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences, 102,* 17245–17250.

Farroni, T., Massaccesi, S., Menon, E., & Johnson, M. H. (2007). Direct gaze modulates face recognition in young infants. *Cognition, 102,* 396–404.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy, 5,* 39–60.

Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Sciences, 8,* 33–39.

Finzi, D., Gomez, J., Nordt, M., Rezai, A. A., Poltoratski, S., & Grill-Spector, K. (2021). Differential spatial computations in ventral and lateral face-selective regions are scaffolded by structural connections. *Nature Communications, 12,* 2278.

Firestone, C., & Scholl, B. (2017). Seeing physics in the blink of an eye. *Journal of Vision, 17,* 203–203.

Firestone, C., & Scholl, B. J. (under review). Seeing stability: Physical understanding is rooted in automatic visual processing.

Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology, 15,* 447–452.

Florey, J., Clifford, C. W. G., Dakin, S., & Mareschal, I. (2016). Spatial limitations in averaging social cues. *Scientific Reports, 6,* 32210.

Forbes, R. J., & Jackson, P. R. (1980). Non-verbal behaviour and the outcome of selection interviews. *Journal of Occupational Psychology, 53,* 65–72.

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition, 117,* 319–331.

Framorando, D., George, N., Kerzel, D., & Burra, N. (2016). Straight gaze facilitates face processing but does not cause involuntary attentional capture. *Visual Cognition, 24,* 381–391.

Freeman, J. B., Stolier, R. M., & Brooks, J. A. (2020). Dynamic interactive theory as a domain-general account of social perception. In *Advances in Experimental Social Psychology* (Vol. 61, pp. 237–287). Academic Press.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review, 5,* 490–495.

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 319–329.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133,* 694–724.

Galfano, G., Dalmaso, M., Marzoli, D., Pavan, G., Coricelli, C., & Castelli, L. (2012). Eye gaze cannot be ignored (but neither can arrows). *Quarterly Journal of Experimental Psychology, 65,* 1895–1910.

Gallup, A. C., Hale, J. J., Sumpter, D. J. T., Garnier, S., Kacelnik, A., Krebs, J. R., & Couzin,

I. D. (2012). Visual attention and the acquisition of information in human crowds. *Proceedings of the National Academy of Sciences, 109,* 7245–7250.

Gallup, G. G., Cummings, W. H., & Nash, R. F. (1972). The experimenter as an independent variable in studies of animal hypnosis in chickens (Gallus gallus). *Animal Behaviour, 20,* 166–169.

Gamer, M., & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 705–715.

Gamer, M., Hecht, H., Seipp, N., & Hiller, W. (2011). Who is looking at me? The cone of gaze widens in social phobia. *Cognition and Emotion, 25,* 756–764.

Gao, T., McCarthy, G., & Scholl, B. J. (2010). The wolfpack effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science, 21,* 1845–1853.

Gao, T., New, J. J., & Scholl, B. J. (under review). Attention to intention: The perceived goals of moving shapes control how they are attended. Manuscript submitted for publication.

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology, 59,* 154–179.

Gao, T., Scholl, B. J., & McCarthy, G. (2012). Dissociating the detection of intentionality from animacy in the right posterior superior temporal sulcus. *Journal of Neuroscience, 32,* 14276–14280.

George, N., & Conty, L. (2008). Facing the gaze of others. *Neurophysiologie Clinique/Clinical Neurophysiology, 38,* 197–207.

George, N., Driver, J., & Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *NeuroImage, 13,* 1102–1112.

Gergely, G., Nádasdy, Z., Csibra, G., & Biró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56,* 165–193.

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *The American Journal of Psychology, 76,* 386.

Gilad, S., Meng, M., & Sinha, P. (2009). Role of ordinal contrast relationships in face encoding. *Proceedings of the National Academy of Sciences, 106,* 5353–5358.

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition, 26,* 651–658.

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science, 20,* 347–353.

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Hughes, H. C., & Cipolli, C. (2013). Processing of invisible social cues. *Consciousness and Cognition, 22,* 765–770.

Gomes, A. (2015). Testimony and other minds. *Erkenntnis, 80,* 173–183.

Gopnik, A., & Meltzoff, A. N. (1997). Words, thoughts, and theories. Cambridge, MA, US: MIT Press.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language, 1,* 158–171.

Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56,* 544–549.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41,* 2261–2271.

Gründl, M., Knoll, S., Eisenmann-Klein, M., & Prantl, L. (2012). The blue-eyes stereotype: Do eye color, pupil diameter, and scleral color affect attractiveness? *Aesthetic Plastic Surgery, 36,* 234–240.

Green, D., & Swets, J. (1966). Signal detection theory and psychophysics. New York, NY: Wiley.

Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 290,* 181–197.

Grossmann, T. (2017). The eyes as windows into other minds: An integrative perspective. *Perspectives on Psychological Science, 12,* 107–121.

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., & Vuilleumier, P. (2012). White-matter connectivity between face-responsive regions in the human brain. *Cerebral Cortex, 22,* 1564–1576.

Guastella, A. J., Mitchell, P. B., & Dadds, M. R. (2008). Oxytocin increases gaze to the eye region of human faces. *Biological Psychiatry, 63,* 3–5.

Guterstam, A., Kean, H. H., Webb, T. W., Kean, F. S., & Graziano, M. S. A. (2019). Implicit model of other people's visual attention as an invisible, force-carrying beam projecting from the eyes. *Proceedings of the National Academy of Sciences, 116,* 328–333.

Haith, M., Bergman, T., & Moore, M. (1977). Eye contact and face scanning in early infancy. *Science, 198,* 853–855.

Hamlin, J. K. (2015). The case for social evaluation in preverbal infants: Gazing toward one's goal drives infants' preferences for Helpers over Hinderers in the hill paradigm. *Frontiers* *in Psychology, 5,* 1563.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450,* 557–559.

Harbort, J., Witthöft, M., Spiegel, J., Nick, K., & Hecht, H. (2013). The widening of the gaze cone in patients with social anxiety disorder and its normalization after CBT. *Behaviour Research and Therapy, 51,* 359–367.

Hare, B. (2002). The domestication of social cognition in dogs. *Science, 298,* 1634–1636.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour, 59,* 771–785.

Harrison, N. A., Singer, T., Rotshtein, P., Dolan, R. J., & Critchley, H. D. (2006). Pupillary contagion: Central mechanisms engaged in sadness processing. *Social Cognitive and Affective Neuroscience, 1,* 5–17.

Hayward, D. A., & Ristic, J. (2013). The uniqueness of social attention revisited: Working memory load interferes with endogenous but not social orienting. *Experimental Brain Research, 231,* 405–414.

Heal, J. (1986). Replication and functionalism. In Butterfield (Ed.), *Language Mind and Logic* (Vol. 14, pp. 135–150). Cambridge, UK: Cambridge University Press.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology, 57,* 243–259.

Hein, G., & Knight, R. T. (2008). Superior temporal sulcus—it's my area: Or is it? *Journal of Cognitive Neuroscience, 20,* 2125–2136.

Heinrichs, M., von Dawans, B., & Domes, G. (2009). Oxytocin, vasopressin, and human social behavior. *Frontiers in Neuroendocrinology, 30,* 548–557.

Hemsley, G. D., & Doob, A. N. (1978). The effect of looking behavior on perceptions of a communicator's credibility. *Journal of Applied Social Psychology, 8,* 136–142.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33,* 98–106.

Hennig, C. W. (1977). Effects of simulated predation on tonic immobility in Anolis carolinensis: The role of eye contact. *Bulletin of the Psychonomic Society, 9,* 239–242.

Hess, E. H. (1965). Attitude and pupil size. *Scientific American, 212,* 46–55.

Hess, E. H. (1975). The role of pupil size in communication. *Scientific American, 233,* 110–119.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science,* *132,* 349–350.

Hicks, R. A., Reaney, T., & Hill, L. (1967). Effects of pupil size and facial angle on preference for photographs of a young woman. *Perceptual and Motor Skills, 24,* 388–390.

Hietanen, J. K. (2018). Affective eye contact: An integrative review. *Frontiers in Psychology, 9,* 1587.

Hietanen, J. K., & Leppänen, J. M. (2003). Does facial expression affect attention orienting by gaze direction cues? *Journal of Experimental Psychology: Human Perception and Performance, 29,* 1228–1243.

Hietanen, J. K., Helminen, T. M., Kiilavuori, H., Kylliäinen, A., Lehtonen, H., & Peltola, M. J. (2018). Your attention makes me smile: Direct gaze elicits affiliative facial expressions. *Biological Psychology, 132,* 1–8.

Hietanen, J. K., Leppänen, J. M., Peltola, M. J., Linna-aho, K., & Ruuhiala, H. J. (2008). Seeing direct and averted gaze activates the approach–avoidance motivational brain systems. *Neuropsychologia, 46,* 2423–2430.

Hietanen, J. K., Nummenmaa, L., Nyman, M. J., Parkkola, R., & Hämäläinen, H. (2006). Automatic attention orienting by social and symbolic cues activates different neural networks: An fMRI study. *NeuroImage, 33,* 406–413.

Hoffman, D. D. (2000). Visual intelligence: How we create what we see. WW Norton & Company.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience, 3,* 80–84.

Hollands, M. A., Patla, A. E., & Vickers, J. N. (2002). mydoubleqLook where you're going!: Gaze behaviour associated with maintaining and changing the direction of locomotion. *Experimental Brain Reserach, 143,* 221–230.

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science, 9,* 131–134.

Hung, S., Nieh, C., & Hsieh, P. (2016). Unconscious processing of facial attractiveness: Invisible attractive faces orient visual attention. *Scientific Reports, 6,* 37117.

Ishikawa, M., Senju, A., & Itakura, S. (2020). Learning process of gaze following: Computational modeling based on reinforcement learning. *Frontiers in Psychology, 11,* 213.

Itakura, S. (1996). An exploratory study of gaze-monitoring in nonhuman primates. *Japanese*

Psychological Research, 38, 174–180.

Itakura, S., & Tanaka, M. (1998). Use of experimenter-given cues during object-choice tasks by chimpanzees (pan troglodytes), an orangutan (pongopygmaeus), and human infants (homosapiens). *Journal of Comparative Psychology, 112,* 119–126.

Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience & Biobehavioral Reviews, 33,* 843–863.

Itier, R. J., Latinus, M., & Taylor, M. J. (2006). Face, eye, and object early processing: What is the face specificity? *Neuroimage, 29,* 667–676.

Janik, S. W., Wellens, A. R., Goldberg, M. L., & Dell'Osso, L. F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual and Motor Skills, 47,* 857–858.

Jasso, H., Triesch, J., Deák, G., & Lewis, J. M. (2012). A unified account of gaze following. IEEE *Transactions on Autonomous Mental Development, 4,* 257–272.

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

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience, 6,* 766–774.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40,* 1–19.

Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences, 4,* 22–28.

Johnson, S. C., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1,* 233–238.

Jones, B. C., DeBruine, L. M., Little, A. C., Conway, C. A., & Feinberg, D. R. (2006). Integrating gaze direction and expression in preferences for attractive faces. *Psychological Science, 17,* 588–591.

Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L. M., Feinberg, D. R., & Tiddeman, B. P. (2010). Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. *Proceedings of the Royal Society B: Biological Sciences, 277,* 617–624.

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature, 504,* 427–431.

Jun, Y. Y., Mareschal, I., Clifford, C. W. G., & Dadds, M. R. (2013). Cone of direct gaze as a marker of social anxiety in males. *Psychiatry Research, 210,* 193–198.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154,* 1583–1585.

Kajimura, S., & Nomura, M. (2016). When we cannot speak: Eye contact disrupts resources available to cognitive control processes during verb generation. *Cognition, 157,* 352–357.

Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2005). Domestic goats, Capra hircus, follow gaze direction and use social cues in an object choice task. *Animal Behaviour, 69,* 11–18.

Kampe, K. K. W., Frith, C. D., & Frith, U. (2003). mydoubleqHey John: Signals conveying communicative intention toward the self activate brain regions associated with mydoubleqmentalizing, regardless of modality. *The Journal of Neuroscience, 23,* 5258–5263.

Kampe, K. K. W., Frith, C. D., Dolan, R. J., & Frith, U. (2001). Reward value of attractiveness and gaze. *Nature, 413,* 589–589.

Kano, F., & Tomonaga, M. (2010). Face scanning in chimpanzees and humans: Continuity and discontinuity. *Animal Behaviour, 79,* 227–235.

Kano, F., Call, J., & Tomonaga, M. (2012). Face and eye scanning in gorillas (Gorilla gorilla), orangutans (Pongo abelii), and humans (Homo sapiens): Unique eye-viewing patterns in humans among hominids. *Journal of Comparative Psychology, 126,* 388–398.

Kano, F., Moore, R., Krupenye, C., Hirata, S., Tomonaga, M., & Call, J. (2018). Human ostensive signals do not enhance gaze following in chimpanzees, but do enhance object-oriented attention. *Animal Cognition, 21,* 715–728.

Kaplan, G., & Rogers, L. J. (2002). Patterns of gazing in orangutans (pongo pygmaeus). *International Journal of Primatology, 23,* 501–52.

Kawai, N. (2011). Attentional shift by eye gaze requires joint attention: Eye gaze cues are unique to shift attention. *Japanese Psychological Research, 53,* 292–301.

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., ... Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring: A PET study. *Brain, 122,* 779–783.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26,* 22–63.

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Sci-*

ence, 330, 932.

King, D., Rowe, A., & Leonards, U. (2011). I trust you; hence I like the things you look at: Gaze cueing and sender trustworthiness influence object evaluation. *Social Cognition, 29,* 476–485.

Kingstone, A., Kachkovski, G., Vasilyev, D., Kuk, M., & Welsh, T. N. (2019). Mental attribution is not sufficient or necessary to trigger attentional orienting to gaze. *Cognition, 189,* 35–40.

Kingstone, A., Smilek, D., Ristic, J., Kelland Friesen, C., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science, 12,* 176–180.

Kingstone, A., Tipper, C., Ristic, J., & Ngan, E. (2004). The eyes have it!: An fMRI investigation. *Brain and Cognition, 55,* 269–271.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin, 100,* 78–100.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature, 387,* 767–768.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: Comparative studies on external morphology of the primate eye. *Journal of Human Evolution, 40,* 419–435.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13,* 1292–1298.

Kuhn, G., & Benson, V. (2007). The influence of eye-gaze and arrow pointing distractor cues on voluntary eye movements. *Perception & Psychophysics, 69,* 966–971.

Kuhn, G., & Kingstone, A. (2009). Look away! Eyes and arrows engage oculomotor responses automatically. *Attention, Perception & Psychophysics, 71,* 314–327.

Kuhn, G., Vacaityte, I., D'Souza, A. D. C., Millett, A. C., & Cole, G. G. (2018). Mental states modulate gaze following, but not automatically. *Cognition, 180,* 1–9.

Kylliäinen, A., & Hietanen, J. K. (2004). Attention orienting by another's gaze direction in children with autism. *Journal of Child Psychology and Psychiatry, 45,* 435–444.

Laeng, B., & Alnæs, D. (2019). Pupillometry. In C. Klein & U. Ettinger (Eds.), Eye Movement Research (pp 449–502). Cham, Switzerland: Springer International Publishing.

Laeng, B., & Sulutvedt, U. (2014). The eye pupil adjusts to imaginary light. *Psychological*

Science, 25, 188–197.

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science, 7,* 18–27.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41,* 3559–3565.

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception, 28,* 1311–1328.

Langton, S. R. H, Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4,* 50–59.

Langton, S. R. H. (2009). Gaze perception: Is seeing influenced by believing? *Current Biology, 19,* R651–R653.

Langton, S. R. H., & Bruce, V. (1999). Reflexive Visual Orienting in Response to the Social Attention of Others. *Visual Cognition, 6,* 541–567.

Langton, S. R. H., & Bruce, V. (2000). You must see the point: Automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 747–757.

Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4,* 50–59.

Lavín, C., San Martín, R., & Rosales Jubal, E. (2014). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience, 7,* 218.

Lavelle, J. S. (2012). Theory-theory and the direct perception of mental states. *Review of Philosophy and Psychology, 3,* 213–230.

Law, A. S., Langton, S. R. H., & Logie, R. H. (2010). Assessing the impact of verbal and visuospatial working memory load on eye-gaze cueing. *Visual Cognition, 18,* 1420–1438.

Lee, D. H., & Anderson, A. K. (2017). Reading what the mind thinks from how the eye sees. *Psychological Science, 28,* 494–503.

Leekam, S., Baron-Cohen, S., Perrett, D., Milders, M., & Brown, S. (1997). Eye-direction detection: A dissociation between geometric and joint attention skills in autism. *British Journal of Developmental Psychology, 15,* 77–95.

Leslie, A. M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology, 6,* 315–324.

Lewis, M. B., & Edmonds, A. J. (2003). Face detection: Mapping human performance. *Per-*

ception, 32, 903–920.

Lobmaier, J. S., & Perrett, D. I. (2011). The world smiles at me: Self-referential positivity bias when interpreting direction of attention. *Cognition and Emotion, 25,* 334–341.

Lobmaier, J. S., Fischer, M. H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology, 53,* 117–122.

Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy: How, when, and where we perceive life in a face. *Psychological Science, 21,* 1854–1862.

Mack, A. (2003). Inattentional blindness: Looking without seeing. *Current Directions in Psychological Science, 12,* 180–184.

Mack, A., & Rock, I. (1998). Inattentional blindness: Perception without attention. In R. D. Wright (Ed.), *Visual Attention* (Vol. 8, pp. 55–76). New York, NY, US: Oxford University Press.

MacLean, E. L., & Hare, B. (2012). Bonobos and chimpanzees infer the target of another's attention. *Animal Behaviour, 83,* 345–353.

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science, 13,* 460–464.

Mansfield, E., Farroni, T., & Johnson, M. (2003). Does gaze perception facilitate overt orienting? *Visual Cognition, 10,* 7–14.

Mar, R. A., Mason, M. F., & Litvack, A. (2012). How daydreaming relates to life satisfaction, loneliness, and social support: The importance of gender and daydream content. *Consciousness and Cognition, 21,* 401–407.

Mareschal, I., Calder, A. J., & Clifford, C. W. G. (2013). Humans have an expectation that gaze is directed toward them. *Current Biology, 23,* 717–721.

Markson, L., & Paterson, K. B. (2009). Effects of gaze-aversion on visual-spatial imagination. *British Journal of Psychology, 100,* 553–563.

Mason, M. F., Hood, B. M., & Macrae, C. N. (2004). Look into my eyes: Gaze direction and person memory. *Memory, 12,* 637–643.

Mason, M. F., Tatkow, E. P., & Macrae, C. N. (2005). The look of love: Gaze shifts and person perception. *Psychological Science, 16,* 236–239.

Mattavelli, G., Romano, D., Young, A. W., & Ricciardelli, P. (in press). The interplay between gaze cueing and facial trait impressions. *Quarterly Journal of Experimental Psychology.*

Matthis, J. S., Yates, J. L., & Hayhoe, M. M. (2018). Gaze and the control of foot placement when walking in natural terrain. *Current Biology, 28,* 1224-1233.

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6,* 255–260.

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex, 9,* 431–444.

McKelvie, S. J. (1976). The role of eyes and mouth in the memory of a face. *The American Journal of Psychology, 89,* 311–323.

Meltzoff, A. N., & Brooks, R. (2008). Self-experience as a mechanism for learning about others: A training study in social cognition. *Developmental Psychology, 44,* 1257–1265.

Meltzoff, A. N., Brooks, R., Shon, A. P., & Rao, R. P. N. (2010). mydoubleqSocial robots are psychological agents for infants: A test of gaze following. *Neural Networks, 23,* 966–972.

Meyerhoff, H. S., Huff, M., & Schwan, S. (2013). Linking perceptual animacy to attention: Evidence from the chasing detection paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1003–1015.

Michotte, A. (1991). The emotions regarded as functional connections. In G. Thinés, A. Costall, & G. Butterworth (Eds.), Michotte's Experimental Phenomenology of Perception (pp. 103–116). Hillsdale, NJ: Erlbaum. [Reprinted from M. Reymert (Ed.). (1950). Feelings and emotions: The Mooseheart symposium (pp. 114–125). New York, NY: McGraw-Hill]

Miles, M. (1983). Vision: The eye of the body and the eye of the mind in Saint Augustine's mydoubleqDe Trinitate and mydoubleqConfessions. *The Journal of Religion, 63,* 125–142.

Milgram, S., Bickman, L., & Berkowitz, L. (1969). Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology, 13,* 79–82.

Miyazaki, Y., Wake, H., Ichihara, S., & Wake, T. (2012). Attentional bias to direct gaze in a dot-probe paradigm. *Perceptual and Motor Skills, 114,* 1007–1022.

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science, 7,* F1–F9.

Moore, C., & Corkum, V. (1994). Social understanding at the end of the first year of life.

Developmental Review, 14, 349–372.

Moors, P., Gayet, S., Hedger, N., Stein, T., Sterzer, P., van Ee, R., … Hesselmann, G. (2019). Three criteria for evaluating high-level processing in continuous flash suppression. *Trends in Cognitive Sciences, 23,* 267–269.

Morgan, E. J., Freeth, M., & Smith, D. T. (2018). Mental state attributions mediate the gaze cueing effect. *Vision, 2,* 11.

Mosconi, M. W., Mack, P. B., McCarthy, G., & Pelphrey, K. A. (2005). Taking an mydoubleqintentional stance on eye-gaze shifts: A functional neuroimaging study of social perception in children. *NeuroImage, 27,* 247–252.

Movellan, J., & Watson, J. (1987). Perception of directional attention. Infant Behavior and Development: Abstracts of the Sixth International Conference of Infant Studies.

Myllyneva, A., & Hietanen, J. K. (2015). There is more to eye contact than meets the eye. *Cognition, 134,* 100–109.

Nakamura, K., & Kawabata, H. (2018). Preferential access to awareness of attractive faces in a breaking continuous flash suppression paradigm. *Consciousness and Cognition, 65,* 71–82.

Nawroth, C., Trincas, E., & Favaro, L. (2017). African penguins follow the gaze direction of conspecifics. *PeerJ, 5,* e3459.

Neal, T. M. S., & Brodsky, S. L. (2008). Expert witness credibility as a function of eye contact behavior and gender. *Criminal Justice and Behavior, 35,* 1515–1526.

Nettle, D., Cronin, K. A., & Bateson, M. (2013). Responses of chimpanzees to cues of conspecific observation. *Animal Behaviour, 86,* 595–602.

Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I. S., & Bateson, M. (2013). The watching eyes effect in the Dictator Game: It's not how much you give, it's being seen to give something. *Evolution and Human Behavior, 34,* 35–40.

New, J. J., & Scholl, B. J. (2008). mydoubleqPerceptual scotomas: A functional account of motion-induced blindness. *Psychological Science, 19,* 653–659.

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences, 104,* 16598–16603.

Nie, Q.-Y., Ding, X., Chen, J., & Conci, M. (2018). Social attention directs working memory maintenance. *Cognition, 171,* 85–94.

Northover, S. B., Pedersen, W. C., Cohen, A. B., & Andrews, P. W. (2017). Artificial surveillance cues do not increase generosity: Two meta-analyses. *Evolution and Human Behavior, 38,* 144–153.

Nuku, P., & Bekkering, H. (2008). Joint attention: Inferring what others perceive (and don't perceive). *Consciousness and Cognition, 17,* 339–349.

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences, 13,* 135–143.

Okumura, Y., Kanakogi, Y., Kanda, T., Ishiguro, H., & Itakura, S. (2013). Infants understand the referential nature of human gaze but not robot gaze. *Journal of Experimental Child Psychology, 116,* 86–95.

Ostrovsky, Y., Andalman, A., & Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science, 17,* 1009–1014.

Otsuka, Y., Mareschal, I., & Clifford, C. W. G. (2016). Testing the dual-route model of perceived gaze direction: Linear combination of eye and head cues. *Journal of Vision, 16,* 8.

Pönkänen, L. M., Peltola, M. J., & Hietanen, J. K. (2011). The observer observed: Frontal EEG asymmetry and autonomic responses differentiate between another person's direct and averted gaze when the face is seen live. *International Journal of Psychophysiology, 82,* 180–187.

Palanica, A., & Itier, R. J. (2011). Searching for a perceived gaze direction using eye tracking. *Journal of Vision, 11,* 1–13.

Palmer, C. J., & Clifford, C. W. G. (2018b). Adaptation to other people's eye gaze reflects habituation of high-level perceptual representations. *Cognition, 180,* 82–90.

Palmer, C. J., Otsuka, Y., & Clifford, C. W. G. (2020). A sparkle in the eye: Illumination cues and lightness constancy in the perception of eye contact. *Cognition, 205,* 104419.

Pargetter, R. (1984). The scientific inference to other minds. *Australasian Journal of Philosophy, 62,* 158–163.

Peacocke, C., & McGinn, C. (1984). Consciousness and other minds. *Proceedings of the Aristotelian Society, Supplementary Volumes, 58,* 97–137.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162,* 8–13.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51,* 195–203.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders, 32,* 249–261.

Pelphrey, K. A., Singerman, J. D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: The influence of context. *Neuropsychologia, 41,* 156–170.

Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science, 15,* 598–603.

Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind.' *Philosophical Transactions of the Royal Society B: Biological Sciences, 362,* 731–744.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences, 335,* 23–30.

Peshek, D., Semmaknejad, N., Hoffman, D., & Foley, P. (2011). Preliminary evidence that the limbal ring influences facial attractiveness. *Evolutionary Psychology, 9,* 137–146.

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences, 109,* E3314–E3323.

Petrican, R., English, T., Gross, J. J., Grady, C., Hai, T., & Moscovitch, M. (2013). Friend or foe? Age moderates time-course specific responsiveness to trustworthiness cues. *The Journals of Gerontology: Series B, 68,* 215–223.

Pfeiffer, U. J., Schilbach, L., Timmermans, B., Kuzmanovic, B., Georgescu, A. L., Bente, G., & Vogeley, K. (2014). Why we interact: On the functional role of the striatum in the subjective experience of social interaction. *NeuroImage, 101,* 124–137.

Phelps, F. G., Doherty-Sneddon, G, & Warnock, H. (2006). Helping children think: Gaze aversion and teaching. *British Journal of Developmental Psychology, 24,* 577–588.

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a third visual pathway specialized for social perception. *Trends in Cognitive Sciences, 25,* 100–110.

Pitcher, D., Pilkington, A., Rauth, L., Baker, C., Kravitz, D. J., & Ungerleider, L. G. (2020).

The human posterior superior temporal sulcus samples visual space differently from other face-selective regions. *Cerebral Cortex, 30,* 778–785.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25.

Povinelli, D. J., & Eddy, T. J. (1996). Factors influencing young chimpanzees' (pan troglodytes) recognition of attention. *Journal of Comparative Psychology, 110,* 336–345.

Povinelli, D. J., & Eddy, T. J. (1997). Specificity of gaze-following in young chimpanzees. *British Journal of Developmental Psychology, 15,* 213–222.

Povinelli, D. J., & Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind & Language, 19,* 1–28.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1,* 515–526.

Price, H. H. (1931). Our knowledge of other minds. *Proceedings of the Aristotelian Society, 32,* 53–78.

Prochazkova, E., & Kret, M. E. (2017). Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neuroscience & Biobehavioral Reviews, 80,* 99–114.

Provine, R. R., Cabrera, M. O., & Nave-Blodgett, J. (2013). Red, yellow, and super-white sclera: Uniquely human cues for healthiness, attractiveness, and age. *Human Nature, 24,* 126–136.

Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex, 9,* 445–458.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience, 18,* 2188–2199.

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22,* 341–365.

Quadflieg, S., Mason, M. F., & Macrae, C. N. (2004). The owl and the pussycat: Gaze cues and visuospatial orienting. *Psychonomic Bulletin & Review, 11,* 826–831.

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21,* 1300–1310.

Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., & Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications, 7:13289,* 1–7.

Riby, D. M., Doherty-Sneddon, G., & Whittle, L. (2012). Face-to-face interference in typical and atypical development. *Developmental Science, 15,* 281–291.

Ricciardelli, P., Baylis, G., & Driver, J. (2000). The positive and negative of human expertise in gaze perception. *Cognition, 77,* B1–B14.

Ricciardelli, P., Betta, E., Pruner, S., & Turatto, M. (2009). Is there a direct link between gaze perception and joint attention behaviours? Effects of gaze contrast polarity on oculomotor behaviour. *Experimental Brain Research, 194,* 347–357.

Ricciardelli, P., Bricolo, E., Aglioti, S., & Chelazzi, L. (2003). My eyes want to look where your eyes are looking: Exploring the tendency to imitate another individual's gaze. *Neuroreport, 13,* 2259–2264.

Rimmele, U., & Lobmaier, J. S. (2012). Stress increases the feeling of being looked at. *Psychoneuroendocrinology, 37,* 292–298.

Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review, 9,* 507–513.

Ristic, J., Mottron, L., Friesen, C. K., Iarocci, G., Burack, J. A., & Kingstone, A. (2005). Eyes are special but not for everyone: The case of autism. *Cognitive Brain Research, 24,* 715–718.

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

Rochat, P., Morgan, R., & Carpenter, M. (1997). Young infants' sensitivity to movement information specifying social causality. *Cognitive Development, 12,* 537–561.

Rock, I. (1983). The logic of perception. Cambridge, MA, US: MIT Press.

Rosati, A. G., & Hare, B. (2009). Looking past the model species: Diversity in gaze-following skills across primates. *Current Opinion in Neurobiology, 19,* 45–51.

Rothkirch, M., Madipakkam, A. R., Rehn, E., & Sterzer, P. (2015). Making eye contact without awareness. *Cognition, 143,* 108–114.

Ruiz, A., Gómez, J. C., Roeder, J. J., & Byrne, R. W. (2009). Gaze following and gaze priming in lemurs. *Animal Cognition, 12,* 427–434.

Russell, B. (1923). Human knowledge: Its scope and limits. Routledge.

Russell, R., Sweda, J. R., Porcheron, A., & Mauger, E. (2014). Sclera color changes with age and is a cue for perceiving age, health, and beauty. *Psychology and Aging, 29,* 626–635.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience, 5,* 631–632.

Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience, 7,* 499–500.

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, Macaca mulatta, know what others can and cannot hear. *Animal Behaviour, 71,* 1175–1181.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology, 16,* 235–239.

Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia, 42,* 1435–1446.

Scheumann, M., & Call, J. (2004). The use of experimenter-given cues by South African fur seals (Arctocephalus pusillus). *Animal Cognition, 7,* 224–230.

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., ... Vogeley, K. (2010). Minds made for sharing: Initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience, 22,* 2702–2715.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80,* 1–46.

Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment? In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social perception: Detection and interpretation of animacy, agency, and intention* (pp. 197-230). Cambridge, MA: MIT Press.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4,* 299–309.

Schulze, L., Renneberg, B., & Lobmaier, J. S. (2013). Gaze perception in social anxiety and social anxiety disorder. *Frontiers in Human Neuroscience, 7,* 872.

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science, 13,* 402–409.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14,* 391–396.

Semyonov, O., Ziv-El, A., Krumhuber, E., Karasik, S., & Aviezer, H. (2019). Beyond shared signals: The role of downward gaze in the stereotypical representation of sad facial expressions.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18,* 668–671.

Senju, A., & Hasegawa, T. (2005). Direct gaze captures visuospatial attention. *Visual Cognition, 12,* 127–144.

Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences, 13,* 127–134.

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition, 12,* 1474–1496.

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others?: A critical test. *Psychological Science, 22,* 878–880.

Senju, A., Tojo, Y., Dairoku, H., & Hasegawa, T. (2004). Reflexive orienting in response to eye gaze and an arrow in children with and without autism. *Journal of Child Psychology and Psychiatry, 45,* 445–458.

Senju, A., Vernetti, A., Ganea, N., Hudry, K., Tucker, L., Charman, T., & Johnson, M. H. (2015). Early social experience affects the development of eye gaze processing. *Current Biology, 25,* 3086–3091.

Senju, A., Yaguchi, K., Tojo, Y., & Hasegawa, T. (2003). Eye contact does not facilitate detection in children with autism. *Cognition, 89,* B43–B51.

Shepherd, S. V., & Platt, M. L. (2007). Spontaneous social orienting and gaze following in ringtailed lemurs (Lemur catta). *Animal Cognition, 11,* 13–20.

Shepherd, S. V., Deaner, R. O., & Platt, M. L. (2006). Social status gates social attention in monkeys. *Current Biology, 16,* R119–R120.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6,* 1317–1322.

Shirama, A. (2012). Stare in the crowd: Frontal face guides overt attention independently of its gaze direction. *Perception, 41,* 447–459.

Simion, C., & Shimojo, S. (2006). Early interactions between orienting, visual sampling and decision making in facial preference. *Vision Research, 46,* 3331–3335.

Sinha, P. (2000). Here's looking at you, kid. *Perception, 29,* 1005–1008.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science, 5,* 679–692.

Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010). Out of mind, out of sight: Eye blinking as indicator and embodiment of mind wandering. *Psychological Science, 21,* 786–789.

Smith, A. D., Hood, B. M., & Hector, K. (2006). Eye remember you two: Gaze direction modulates face recognition in a developmental study. *Developmental Science, 9,* 465–472.

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16,* 184–189.

Sodian, B., & Thoermer, C. (2004). Infants' understanding of looking, pointing, and reaching as cues to goal-directed action. *Journal of Cognition and Development, 5,* 289–316.

Song, X., & Wang, X. (2012). Mind wandering in chinese daily lives—an experience sampling study. *PLoS ONE, 7,* e44423.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology, 45,* 1794–1798.

Spezio, M. L., Huang, P.-Y. S., Castelli, F., & Adolphs, R. (2007). Amygdala damage impairs eye contact during conversations with real people. *Journal of Neuroscience, 27,* 3994–3997.

Stein, T. (2019). The breaking continuous flash suppression paradigm: Review, evaluation, and outlook. In G. Hesselmann (Ed.), Transitions between Consciousness and Unconsciousness (pp. 1–38). New York: Routledge.

Stein, T., Senju, A., Peelen, M. V., & Sterzer, P. (2011). Eye contact facilitates awareness of faces during interocular suppression. *Cognition, 119,* 307–311.

Sun, Z., Yu, W., Zhou, J., & Shen, M. (2017). Perceiving crowd attention: Gaze following in human crowds with conflicting cues. *Attention, Perception, & Psychophysics, 79,* 1039–1049.

Symons, L. A., Lee, K., Cedrone, C. C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *The Journal of General Psychology, 131,* 451–469.

Takahashi, K., & Watanabe, K. (2013). Gaze cueing by pareidolia faces. *I-Perception, 4,* 490–492.

Takahashi, K., Meilinger, T., Watanabe, K., & Bülthoff, H. H. (2013). Psychological influences on distance estimation in a virtual reality environment. *Frontiers in Human Neuroscience, 7,* 580.

Tecwyn, E. C., Bechlivanidis, C., Lagnado, D. A., Hoerl, C., Lorimer, S., Blakey, E., ... Buehner, M. J. (2020). Causality influences children's and adults' experience of temporal order. *Developmental Psychology, 56,* 739–755.

Teufel, C., Alexis, D. M., Clayton, N. S., & Davis, G. (2010). Mental-state attribution drives rapid, reflexive gaze following. *Attention, Perception, & Psychophysics, 72,* 695–705.

Teufel, C., Alexis, D. M., Todd, H., Lawrance-Owen, A. J., Clayton, N. S., & Davis, G. (2009). Social cognition modulates the sensory coding of observed gaze direction. *Current Biology, 19,* 1274–1277.

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: Attributed mental states influence perception. *Trends in Cognitive Sciences, 14,* 376–382.

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review, 9,* 314–318.

Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. Animal Behaviour.

Tomasello, M., Hare, B., & Agnetta, B. (1999). Chimpanzees, Pan troglodytes, follow gaze direction geometrically. *Animal Behaviour, 58,* 769–777.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution, 52,* 314–320.

Tombs, S., & Silverman, I. (2004). Pupillometry: A sexual selection approach. *Evolution and Human Behavior, 25,* 221–228.

Triesch, J., Teuscher, C., Deak, G. O., & Carlson, E. (2006). Gaze following: Why (not) learn it? *Developmental Science, 9,* 125–147.

Tschudin, A., Call, J., Dunbar, R. I. M., Harris, G., & van der Elst, C. (2001). Comprehension of signs by dolphins (Tursiops truncatus). *Journal of Comparative Psychology, 115,* 100–105.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8,* 1096–1101.

U.S. Bureau of Labor Statistics. (2019). American time use survey (2019). (Version v5) [Data set]. Inter-University Consortium for Political and Social Research.

Ullman, S. (1984). Visual routines. *Cognition, 18,* 97–159.

Ungerleider, L. G., & Desimone, R. (1986). Cortical connections of visual area MT in the macaque. *Journal of Comparative Neurology, 248,* 190–222.

van Buren, B., & Scholl, B. (2018). The 'blindfold test' for deciding whether an effect reflects visual processing or higher-level judgment. *Journal of Vision, 18,* 56–56.

van Buren, B., & Scholl, B. J. (2017). Minds in motion in memory: Enhanced spatial memory driven by the perceived animacy of simple shapes. *Cognition, 163,* 87–92.

van Buren, B., & Scholl, B. J. (under review). The mydoubleqblindfold test for deciding whether an effect reflects visual processing or higher-level judgment.

van Buren, B., Gao, T., & Scholl, B. J. (2017). What are the underlying units of perceived animacy? Chasing detection is intrinsically object-based. *Psychonomic Bulletin & Review, 24,* 1604–1610.

van Buren, B., Uddenberg, S., & Scholl, B. J. (2016). The automaticity of perceiving animacy: Goal-directed motion in simple shapes influences visuomotor behavior even when task-irrelevant. *Psychonomic Bulletin & Review, 23,* 797–802.

van Hooff, J. (1967). The facial displays of catarrhine monkeys and apes. In P. D. Morris (Ed.), Primate Ethology (pp. 7–68.). Routledge.

Vestner, T., Tipper, S. P., Hartley, T., Over, H., & Rueschemeyer, S.-A. (2019). Bound together: Social binding leads to faster processing, spatial distortion, and enhanced memory of interacting partners. *Journal of Experimental Psychology: General, 148,* 1251–1268.

Vine, I. (1971). Judgement of direction of gaze: An interpretation of discrepant results. *British Journal of Social and Clinical Psychology, 10,* 320–331.

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes. *Cognitive Science, 28,* 289–301.

von Grünau, M., & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception, 24,* 1297–1313.

Vuilleumier, P. (2002). Perceived gaze direction in faces and spatial attention: A study in patients with parietal damage and unilateral neglect. *Neuropsychologia, 40,* 1013–1026.

Vuilleumier, P., George, N., Lister, V., Armony, J., & Driver, J. (2005). Effects of perceived mutual gaze and gender on face processing and recognition memory. *Visual Cognition, 12,* 85–101.

Wagenbreth, C., Rieger, J., Heinze, H.-J., & Zaehle, T. (2014). Seeing emotions in the eyes—inverse priming effects induced by eyes expressing mental states. *Frontiers in Psychology, 5,* 1039.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception, 6,* 313–326.

Wang, J. J., & Apperly, I. A. (2017). Just one look: Direct gaze briefly disrupts visual working memory. *Psychonomic Bulletin & Review, 24,* 393–399.

Wang, Y., Newport, R., & Hamilton, A. F. de C. (2011). Eye contact enhances mimicry of intransitive hand movements. *Biology Letters, 7,* 7–10.

Watanabe, S., Kakigi, R., Miki, K., & Puce, A. (2006). Human MT/V5 activity on viewing eye gaze changes in others: A magnetoencephalographic study. *Brain Research, 1092,* 152–160.

West, R. W. (2013). The effect of head turn and illumination on the perceived direction of gaze. *Perception, 42,* 495–507.

Wheeler, R. W., Baron, J. C., Michell, S., & Ginsburg, H. J. (1979). Eye contact and the perception of intelligence. *Bulletin of the Psychonomic Society, 13,* 101–102.

Wicker, B., Michel, F., Henaff, M.-A., & Decety, J. (1998). Brain regions involved in the perception of gaze: A PET study. *NeuroImage, 8,* 221–227.

Wicker, B., Perrett, D. I., Baron-Cohen, S., & Decety, J. (2003). Being the target of another's emotion: A PET study. *Neuropsychologia, 41,* 139–146.

Wiese, E., Buzzell, G. A., Abubshait, A., & Beatty, P. J. (2018). Seeing minds in others: Mind perception modulates low-level social-cognitive performance and relates to ventromedial prefrontal structures. *Cognitive, Affective, & Behavioral Neuroscience, 18,* 837–856.

Wiese, E., Wykowska, A., Zwickel, J., & Müller, H. J. (2012). I see what you mean: How attentional selection is shaped by ascribing intentions to others. *PLoS ONE, 7,* e45391.

Wilkinson, N., Paikan, A., Gredebäck, G., Rea, F., & Metta, G. (2014). Staring us in the face? An embodied theory of innate face preference. *Developmental Science, 17,* 809–825.

Williams, C. C., & Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory & Cognition, 35,* 1977–1985.

Wohltjen, S., & Wheatley, T. (2020). Eye contact is a marker of alignment in conversation. *Collective Intelligence.*

Wollaston, W. H. (1824). On the apparent direction of eyes in a portrait. *Philosophical Trans-*

actions of the Royal Society of London, 114, 247–256.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69,* 1–34.

Yarbus, A. L. (1967). Eye movements and vision. Boston, MA: Springer US.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145.

Yokoyama, T., Sakai, H., Noguchi, Y., & Kita, S. (2014). Perception of direct gaze does not require focus of attention. *Scientific Reports, 4,* 3858.

Yorzinski, J. L., & Miller, J. (2020). Sclera color enhances gaze perception in humans. *PLoS ONE, 15,* e0228275.

Yorzinski, J. L., Harbourne, A., & Thompson, W. (2021). Sclera color in humans facilitates gaze perception during daytime and nighttime. *PLoS ONE, 16,* e0249137.

Yoshizaki, K., & Kato, K. (2011). Contrast polarity of eyes modulates gaze-cueing effect: Contrast polarity of eyes modulates gaze-cueing effect. *Japanese Psychological Research, 53,* 333–340.

This thesis was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. A template that can be used to format a PhD dissertation with this look & feel has been released under the permissive AGPL license, and can be found online at github.com/suchow/Dissertate or from its lead author, Jordan Suchow, at suchow@post.harvard.edu.