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

Clara Colombatto1 · Benjamin van Buren1 · Brian J. Scholl1

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? Seeing 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.

Keywords Eye contact · Working memory · Intentionality · Mind contact

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 influence the processing of other (eye-less) objects in a scene.

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 (Buchanan...
et 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).

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. 1. In particular, we followed the procedure of Wang and Apperly (2017) exactly, but substituted direct and averted mouth stimuli (as in Fig. 2b) for their direct and averted gaze stimuli (as in Fig. 2a). Would this alternate means of conveying directed attention still impair visual working memory for the other properties of objects in the scene?

**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.

**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. 1b), or oriented 45°, 135°, 225°, or 315° within the image plane (for mouths directed away from the observer, as in Fig. 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

Fig. 1 Examples of the mouth stimuli used in Experiments 1 and 2

(a) Eye Gaze
Wang & Apperly, 2017

(b) Mouths
Experiment 1

(c) Control
Experiment 2

Directed at You

Directed at Shapes

Change Detection (d')

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.
any given display. The mouths in each trial either faced the observer (as in the top panel of Fig. 2b; “Directed-at-You”) or faced their respective shapes (as in the middle panel of Fig. 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.

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. 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. 2a).1 Thus, the impairment of visual working memory by direct gaze seems not to require gaze, per se, as long as directed attention and intentions are depicted in other ways.

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. 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.

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, ηp²=.24), a main effect of change type (with changes to shapes detected less accurately than changes to mouths; F(1, 15)=42.79, p<.001, ηp²=.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, ηp²=.89), but no significant interactions (all Fs < 2).
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).

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. 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) x 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, ηp2=.02), no effect of stimulus direction (F(1, 30)=1.55, p=.222, ηp2=.05), and – most importantly – a highly reliable interaction between these factors (F(1, 30)=6.68, p=.015, ηp2=.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.

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.”

Author Note For helpful conversations and/or comments on earlier drafts, we thank Jessica Wang and the members of the Yale Perception & Cognition Laboratory. This project was funded by ONR MURI #N00014-16-1-2007 awarded to BJS, and by a National Science Foundation Graduate Research Fellowship awarded to BvB.

References


