Journal of Experimental Psychology:
Human Perception and Performance
2000, Vol. 26, No. 2, 747–757

0096-1523/00/$5.00 DOI: 10.1037//0096-1523.26.2.747

You *Must* See the Point: Automatic Processing of Cues to the Direction of Social Attention

Stephen R. H. Langton and Vicki Bruce
University of Stirling

Four experiments explored the processing of pointing gestures comprising hand and combined head and gaze cues to direction. The cross-modal interference effect exerted by pointing hand gestures on the processing of spoken directional words, first noted by S. R. H. Langton, C. O'Malley, and V. Bruce (1996), was found to be moderated by the orientation of the gesturer's head–gaze (Experiment 1). Hand and head cues also produced bidirectional interference effects in a within-modalities version of the task (Experiment 2). These findings suggest that both head–gaze and hand cues to direction are processed automatically and in parallel up to a stage in processing where a directional decision is computed. In support of this model, head–gaze cues produced no influence on nondirectional decisions to social emblematic gestures in Experiment 3 but exerted significant interference effects on directional responses to arrows in Experiment 4. It is suggested that the automatic analysis of head, gaze, and pointing gestures occurs because these directional signals are processed as cues to the direction of another individual's social attention.

It has long been realized that spoken language is not the exclusive communication medium available to the human species. In addition to the voice, people use a variety of facial gestures in most social situations. For instance, facial expressions convey a variety of emotions and interpersonal attitudes (see Argyle, 1988); movements of the lips, teeth, and tongue assist people's interpretations of what is said to them (McGurk & MacDonald, 1976); and gaze appears to be useful in regulating turn taking in conversation, in expressing certain attitudes such as liking or loving, and in exercising social control (see Kleinke, 1986, for a review).

Along with facial expression, gaze, and lip configurations, spontaneous speech is also accompanied by a myriad of gestures. During speech, the hands are in a constant state of flux, either making rhythmical movements of varying amplitude or describing particular shapes and patterns in the air. However, despite their ubiquity, there is little agreement about the functions performed by these speech-related gestures or about the processes by which they are generated and understood (for reviews, see Kendon, 1994; Rimé & Schiaratura, 1991). In particular, except for a handful of recent studies (Langton, O'Malley, & Bruce, 1996; Thompson & Massaro, 1986, 1994), gesture *comprehension* remains a "neglected field in cognitive psychology" (Feyereisen, 1991, p. 57). Accordingly, in this article we focus on the comprehension of a particular type of cospeech gesture: the deictic or pointing gesture.

The popular idea that gestures and other aspects of nonverbal behavior form part of a separate "body language" carries with it the common assumption that a speaker's gestural behavior plays a critical role in communication. However, this view is not universally accepted. Some have argued that gestures are produced for the benefit of the speaker, playing an instrumental role in maintaining the fluency of speech by facilitating the retrieval of items from lexical memory (e.g., Morrel-Samuels & Krauss, 1992; Rimé & Schiaratura, 1991). It follows from this view that the information conveyed by gesture is largely redundant with speech. Accordingly, under normal circumstances, listeners are not expected to attend to or to process a speaker's gestural activity (Rimé & Schiaratura, 1991). On the other hand, researchers such as Kendon (1986, 1994) and McNeill (1985, 1987, 1989) are enthusiastic supporters of the idea that gestures play a critical role in the communicative process. They argued that gestural and verbal behaviors represent different aspects of the underlying meaning that a speaker is striving to express. It follows that a listener must process both the gestural and the verbal components of the utterance, combining this information at some point in processing to provide an integrated representation of the speaker's intended meaning (Langton et al., 1996; McNeill, Cassell, & McCullough, 1994).

A number of studies have yielded results that are consistent with the suggestion that listeners do indeed process gestural information in comprehension (see Kendon, 1994, for a review). These studies have shown that listeners' understanding of verbal utterances are influenced by speakers' accompanying gestural performances. Langton et al. (1996) explored this issue using a variant of the Stroop interference paradigm (Stroop, 1935).

The Stroop paradigm has proved useful to researchers in a variety of disciplines as a tool with which to demonstrate the automatic processing of a to-be-ignored or irrelevant aspect

Stephen R. H. Langton and Vicki Bruce, Department of Psychology, University of Stirling, Stirling, Scotland.

We thank Carlo Umiltà and three anonymous reviewers for their helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Stephen R. H. Langton, Department of Psychology, University of Stirling, Stirling, FK9 4LA, Scotland. Electronic mail may be sent to srhl1@stirling.ac.uk.

of a stimulus (see MacLeod, 1991, for a review). For instance, in the original version of the Stroop color word task, participants were slower to name the color of the ink in which an incongruent color word was printed (e.g., RED in blue ink) relative to a control condition of naming color squares. This finding clearly indicates that the to-be-ignored dimension (the color word) nevertheless receives some form of analysis by the perceptual or cognitive system despite participants' intentions. One interpretation of this kind of "breakdown" in selective attention is that the component dimensions under scrutiny interact or exchange information in some way and at some stage in information processing. From this perspective, the observation of an interference effect therefore entails two things: that the to-be-ignored dimension has been processed automatically and that information from this dimension has interacted with information encoded from the target dimension.

Langton et al. (1996) made pointing (deictic) gestures conflict with spoken directional words and asked participants to respond to the direction of the gesture in one block of trials and the identity of the spoken word in another block of trials. Langton et al. showed that vocal or manual responses to spoken words were influenced by to-be-ignored gestures and, reciprocally, that responses to pointing gestures were influenced by spoken words. Thus, participants' reaction times (RTs) to the word *up* were slowed if they simultaneously saw a person pointing downward compared with upward. Similarly, when asked to respond to an upward pointing gesture, RTs were slower when participants heard a voice saying "down" compared with "up." These findings clearly indicate that participants are simply unable to ignore deictic gestures when interpreting speech. More specifically, Langton et al. suggested that these interference effects arise because information from deictic gestures and speech is integrated at some point in the comprehension process.

However, the visual gesture stimuli used by Langton et al. (1996) did not simply consist of extended arms and fingers signaling the direction of the deictic gesture. Strong directional cues were also provided by the orientation of the gesturer's head and the direction of his or her gaze. Thus, it is possible that the interference effects exerted by deictic gestures on spoken words are the result of the automatic processing of any or all of these directional signals.

In the experiments reported here we again used the Stroop-type interference paradigm to examine whether head–gaze cues would be processed automatically along with the directional signals provided by the hands and to investigate how all of these cues might contribute to the cross-modal interference effects exerted by deictic gestures on responses to spoken words. In Experiment 1 we examined whether the interference effect exerted by to-be-ignored hand gestures on spoken words could be moderated by the orientation of the gesturer's head. In Experiment 2 we used an intramodal version of the task to explore whether head orientation could moderate the processing of hand gestures and vice versa. Finally, in Experiments 3 and 4 we tested the hypothesis that the various directional cues, including gestures, head–gaze signals, and arrows, would exert their effects on a stage or on several stages of processing where a directional decision is computed.

Experiment 1

In Experiment 1 we examined whether head–gaze cues would moderate the interference exerted by to-be-ignored pointing gestures in a cross-modal interference task similar to that used by Langton et al. (1996). On each trial, participants were presented with either the word *up* or the word *down* spoken by a male voice. Each word was presented together with a digitized photograph of a male individual making a pointing gesture that could be congruent or incongruent with the meaning of the spoken word. In addition, the gesturer's head could be oriented in the same direction as his gesture, in the opposite direction as his gesture, or straight ahead. Participants were asked to make speeded keypress responses contingent on the meaning of the spoken word and to completely ignore the visually presented information. Following Langton et al., we expected that agreeing head and gesture cues would exert large interference effects on responses to spoken words. However, if the orientation of the head contributes to the interfering effect of the gesture on responses to spoken words, then this congruity effect might be expected to interact with the agreement of the head and gesture. More specifically, the effect of gestures on spoken words should be reduced when the head is angled straight ahead or in the opposite direction to that of the gesture.

Method

Participants. Participants were 18 undergraduates recruited through advertisements. All had normal hearing and normal or corrected-to-normal vision.

Materials and apparatus. The visual stimuli consisted of four digitized images of a male individual who had been instructed to point either upward or downward while orienting his head in the same direction as the pointing gesture, in the opposite direction to the pointing gesture, or straight ahead. Examples of these images are presented in Figure 1. These stimuli subtended approximately 11° of vertical visual angle and 14° of horizontal visual angle and were viewed by participants seated approximately 0.7 m from a 14-in. (35.56-cm) color monitor.

In addition to the gesture stimuli, the visual materials also contained a large question mark that subtended approximately 3° of

Figure 1. Reproductions of the digitized stimuli used in Experiment 1. These images show congruent and incongruent head and gesture cues. Neutral head stimuli are not shown.

vertical visual angle. The verbal stimuli were recorded using Hypercard audio software and edited using SoundEdit software on a Macintosh IIci. Two spoken words ("up" and "down") were recorded and edited to be approximately the same length (0.8 s). These auditory stimuli were combined with each of the visual images to yield the 12 test stimuli that were presented using the SuperLab software on a Macintosh Performa 630.

Design. The audiovisual stimuli were presented in a within-subjects design with two variables: gesture congruity and head agreement. Gesture congruity (congruent or incongruent) was determined by the relationship between the pointing gesture and the spoken word. Head agreement (agree, disagree, or neutral) was determined by the relationship between the orientation of the head and the direction of the gesture.

Procedure. Each trial began with the simultaneous presentation of a visual gesture stimulus and a spoken word. Participants were asked to respond to the verbal information as quickly and accurately as possible by depressing one of two vertically arranged keys on the keypad area of the keyboard (8 for "up" and 2 for "down" responses, respectively). Participants were asked to operate each key with a separate hand and were free to decide which key to operate with which hand. The onset of the auditory stimulus activated the timer, which was stopped when the participant pressed either of the response keys. This response also terminated the presentation of the visual stimulus. The intertrial interval was set at 1,000 ms following the execution of a response. Participants were also instructed to depress the space bar in response to a large question mark that occasionally appeared on the screen. Seven of these question marks appeared in each response block. These trials were included to ensure that the participants actually watched the screen.

Participants completed a set of 14 practice trials on the task. These comprised 2 question mark trials and one of each of the 12 test stimuli. Following the practice block, we gave participants two blocks of 79 trials (72 test trials plus 7 question mark trials). Each of the 12 test stimuli was presented six times in each block, resulting in a total of 24 trials per condition. Four practice trials were also presented immediately before the second experimental block. Trials were randomized within all blocks.

Results

In this and all other experiments reported here, we removed outliers from individual participant's scores by eliminating RTs greater than 2 *SD*s from each cell mean. The resulting mean correct RT scores and the percentage of errors recorded in each condition are reported in Table 1. Overall, RTs were slower when the gesture and voice were incongruent. However, the difference between incongruent and congruent RTs (the congruity effect) was largest when the head direction was in agreement with the direction of the gesture, compared with conditions with neutral head orientation or in which the head direction was not in agreement with the direction of the gesture.

A 2 (gesture congruity) × 3 (head agreement) analysis of variance (ANOVA) performed on the RT data was consistent with these observations. Performance was affected by the congruity of the gesture with the voice stimuli, $F(1, 17) = 17.88$, $p < .01$, but not by the agreement of the head with the gesture ($p = .79$). The main effect of gesture congruity was qualified by an interaction with head agreement, $F(2, 34) = 4.21$, $p < .05$. The effect of congruity was strong when the head direction agreed with the gesture (61 ms) but weaker when the head direction was either neutral (34 ms) or disagreed with the gesture (28 ms). However, a simple main effects analysis revealed that the effect of gesture congruity was significant at all levels of head agreement (ps $< .05$).

To further examine the interaction between head agreement and gesture congruity, we conducted an additional one-way ANOVA to compare congruity effects in the three head agreement conditions. Congruity scores were calculated by subtracting each participant's mean RT for congruent stimuli from his or her mean RT for incongruent trials in each of the head agreement conditions. The one-way ANOVA revealed a main effect of head agreement, $F(2, 34) = 4.28$, $p < .05$. Post hoc Newman–Keuls tests ($\alpha = .05$) indicated that the gesture congruity effect in the head-agree condition was significantly larger than those effects in both the disagree and neutral conditions, which did not differ from one another.

In Table 1, the error scores generally mirror the RT data; the overall mean error score was only 1.37%. The correlation between overall mean RTs and mean error rates was .77, suggesting no evidence of a trade-off between speed and accuracy that might compromise interpretation of the RT data. Because of the low rate of errors, we did not analyze these data further.

Table 1
Mean Reaction Times (RTs; in Milliseconds) and Percentage of Errors in Each Condition of Experiment 1

Head agreement	Congruent gesture		Incongruent gesture		Congruity effect RT
	RT	% of Errors	RT	% of Errors	
Agree	466	1.11	527	2.00	61
Disagree	482	0.44	510	1.33	28
Neutral	482	1.11	516	2.22	34
M	477	0.89	518	1.85	41

Discussion

The results of this experiment replicate those of Langton et al. (1996), who obtained interference effects from irrelevant pointing gestures when participants were asked to make speeded keypress responses to spoken directional words. Pointing gestures produced reliable congruity effects in all conditions of the experiment. However, the congruity effect observed when the head and gesture were oriented in the same direction (as in Langton et al., 1996) was significantly reduced when the head was in a neutral position or oriented in the opposite direction to the gesture. Clearly, head direction does indeed modulate the interfering effect of irrelevant gestures. Both pointing gestures and head orientation therefore receive some kind of obligatory processing by the cognitive system, even when attention is directed to stimuli appearing in a completely separate modality.

Although the findings of Experiment 1 indicate that both head orientation and gesture were able to influence the processing of spoken words, it was not clear how these stimuli exerted their effects. Logically, there seem to be two possibilities. Information encoded from gesture and head cues may be combined before some integrated representation exerting an influence on the processing of spoken words. Alternatively, gesture and head cues may be processed in parallel, exerting independent effects on the processing of these directional words.

To tease apart these alternatives, we reanalyzed the data from Experiment 1 to examine whether the effect of the congruity of the gesture with the direction words would interact with the congruity of the head orientation with these words or whether these effects would be additive. To achieve this, we replaced the head agreement variable with a head congruity variable, which represented the congruent–incongruent relationship between the head orientation and the directional word, rather than the relationship between the head and the gesture. The RTs for each participant were rescored accordingly and the resultant means entered into a 2 (gesture congruity) × 3 (head congruity) ANOVA. This analysis yielded main effects of gesture congruity, $F(1, 17) = 17.88$, $p < .01$, and head congruity, $F(2, 34) = 4.65$, $p < .05$, and, critically, no interaction between these variables, $F(2, 34) = 0.55$, $p = .58$. Gesture direction and head orientation therefore exerted independent effects on responses to spoken directional words. Therefore, rather than reflecting the influence of a combined gesture–head signal on spoken words, the results are more consistent with a model in which gesture and head orientation are evaluated in parallel, with the outputs of each system influencing the processing of spoken directional words at some stage of processing.

Some evidence from developmental psychology has indicated that the comprehension of gaze may indeed be independent of the comprehension of pointing. Although infants seem to respond to gaze cues from as young as 3 months of age (Hood, Willen, & Driver, 1998) and to head–gaze cues from 6 months of age (e.g., Butterworth & Jarrett, 1991; however, see Corkum & Moore, 1995), comprehension of manual pointing does not seem to emerge until about 12 months of age (Butterworth, cited in Butterworth, 1995). Butterworth argued that, unlike the comprehension of gaze, understanding pointing gestures requires additional cognitive developmental changes—notably the comprehension of signs—that are not available to the developing child until around 10–12 months of age. The comprehension of manual pointing may therefore emerge much later than the understanding of the more "social" cues and may be underpinned by a more general ability to appreciate symbolic meaning.

Langton et al. (1996) obtained symmetrical interference effects between pointing gestures and spoken words. That is, responses to spoken words were influenced by to-be-ignored gestures and, reciprocally, responses to pointing gestures were influenced by irrelevant directional words. They suggested that this pattern of results was best accounted for by a model in which gesture and words were identified in parallel with the outputs of each system somehow combined before a directional response could be determined, programmed, and executed. The findings of the present experiment suggest that the processing of these gestures should be further broken down into the parallel processing of head–gaze orientation and the actual pointing gesture itself. On this basis, one would expect to observe symmetrical interference effects between head and gesture signals when placed into conflict in an intramodal version of the interference task. We examined this hypothesis in Experiment 2.

Experiment 2

In Experiment 2, we presented participants with the agreeing and disagreeing head–gesture stimuli used in Experiment 1. We asked them to make manual keypress responses that were contingent on the direction of the gesture in one block of trials and on the orientation of the head in a separate block of trials. If head and gesture cues are processed automatically and in parallel, then a symmetrical pattern of effects might be expected. Responses to pointing gestures should be influenced by to-be-ignored head cues, and, conversely, responses to head cues should be influenced similarly by irrelevant pointing gestures.

Method

Participants. The participants were 20 undergraduates recruited through advertisements. All had normal or corrected-to-normal vision.

Apparatus and materials. The visual stimuli consisted of four digitized images of a male individual who had been instructed to point either upward or downward while orienting his head in the same direction as the pointing gesture or in the opposite direction to the pointing gesture. Examples of these images are presented in Figure 1. These stimuli subtended approximately 11° of vertical visual angle and 14° of horizontal visual angle, and they were viewed by participants seated approximately 0.7 m from a 14-in. (35.56-cm) color monitor.

Design. The materials were tested in a within-subjects design with three variables: response dimension (head or gesture), congruity (congruent or incongruent), and target direction (up or down).

Procedure. On each trial, participants were asked to respond as quickly and accurately as possible to either the gesture or head direction of the visual stimulus that appeared in the center of the screen. Participants executed their responses by depressing one of two vertically arranged keys on the keypad area of the keyboard (8 for up and 2 for down responses, respectively). They were asked to operate each key with a separate hand and were free to decide which key to operate with which hand. Depression of either key terminated the presentation of the visual stimulus and stopped the timer. After the response, the screen remained blank for 1,000 ms before the start of the next trial.

Participants completed 160 experimental trials: 20 trials in each of the eight cells of the experimental design. These trials were blocked by the response variable and were presented in a random order within each block. Thus, in one block of trials participants responded to the orientation of the head and were asked to ignore the gesture; in a second block of trials they responded to the gesture direction and were asked to ignore the orientation of the head. The order in which these blocks were presented was alternated between successive participants. A set of 16 practice trials was presented before each block that comprised two repetitions of each of the eight experimental stimuli. Both RTs and the proportion of errors were recorded as dependent variables in this experiment.

Results

The interparticipant mean correct RTs and percentage of errors recorded in each condition of the experiment are shown in Table 2. Inspection of this table reveals that, overall, participants' responses were slower and less accurate in the incongruent than in the congruent conditions of the experiment. However, this difference was more marked when head direction was the relevant response dimension.

A 2 (response dimension) × 2 (congruity) × 2 (target direction) ANOVA conducted on participants' mean RTs supported the aforementioned observations. Performance was equivalent for up (474 ms) compared with down (467 ms) targets ($p = .24$). However, there were main effects of both response dimension, $F(1, 19) = 34.28$, $p < .001$, and congruity, $F(1, 19) = 34.41$, $p < .001$, that were qualified by a significant interaction between these variables, $F(1, 19) = 5.23$, $p < .05$. Simple main effects analysis revealed that the effects of congruity were significant for both response dimensions ($ps < .05$). The cause of this interaction therefore appeared to be the relatively strong effect of congruity for responses to head direction (30 ms) rather than the weaker effect of congruity on responses to gesture direction (14 ms).[1]

Error scores closely mirrored the RT data. Indeed, the correlation between the overall mean RTs and mean error rates in each condition was .73, which shows no evidence of a trade-off between speed and accuracy. Because of the low rate of errors (the overall mean error rate was 1.41%), we performed no further analyses on these data.

Table 2
Mean Reaction Times (RTs; in Milliseconds) and Percentage of Errors for Responses to Head and Gesture Direction in Each Condition of Experiment 2

	Down targets		Up targets		Overall mean	
Congruity	RT	% of Errors	RT	% of Errors	RT	% of Errors
Responses to head						
Congruent	486	0.50	496	1.25	491	0.88
Incongruent	521	3.00	520	2.50	521	2.75
M	504	1.75	508	1.88	506	1.82
Responses to gesture						
Congruent	423	0.50	435	1.25	429	0.88
Incongruent	441	1.00	445	1.25	443	1.13
M	432	0.75	440	1.25	436	1.00

Discussion

The results of Experiment 1 show that participants were unable to ignore gesture and head cues to direction even when responding to verbal stimuli. The findings of Experiment 2 also suggest that head and gesture cues are processed automatically by observers. There were significant effects of congruity for responses to both gesture and head–gaze direction stimuli. Thus, to-be-ignored head–gaze directions influenced the participants' responses to the direction of a pointing gesture and, reciprocally, irrelevant pointing gestures influenced speeded judgments of head–gaze direction.

However, despite being bidirectional, the interference effects were not exactly symmetrical; responses to head orientation were affected more by irrelevant gestures than vice versa. One possibility is that this asymmetry may have been the result of mismatching the discriminabilities of the two dimensions of the test stimuli. Melara and Mounts (1993) found that the direction of the interference effects between interacting dimensions was a function of their relative baseline discriminabilities. In particular, the more discriminable dimension (i.e., the dimension with greater trial-to-trial variation) caused greater disruption of classifications of the less discriminable dimension than vice versa. One could argue that in the present experiments the perceptual variation between up and down gestures was greater than that between up and down head angles. Indeed, the fact that participants' responses to gesture stimuli were significantly faster than those to head stimuli (436 vs. 506 ms) provides some evidence for this position. Thus, the more discriminable gesture dimension will exert a greater influence on the less discriminable head orientation dimension than vice versa, resulting in the asymmetrical pattern of interference observed in Experiment 2. However, despite being disadvantaged by relatively poor discriminability, irrelevant head direction cues still produced significant interference effects on classification of the more discriminable pointing gestures.

The findings of this experiment add to those of other studies conducted in our laboratory in which a variety of directional signals produced symmetrical or bidirectional interference effects when placed into conflict in a Stroop-type task. Along with pointing gestures and spoken words, symmetrical interference effects have been obtained between gestures and written directional words, between directional arrows and spoken words, and between head orientation and gaze direction (Langton, in press; Langton et al., 1996). This pattern of results is perhaps best accounted for by a model in which signals such as head orientation, gaze direction, pointing gestures, spoken words, and so on are processed in parallel by several separate systems. These systems feed their outputs into a processing stage, or processing stages in which the directional information automatically encoded from the cues is somehow pooled before a response can be selected and executed. Thus, interference effects will emerge whenever a directional decision has to be made to a directional signal when conflicting information is available from any other copresent cue.

According to this model, it is the nature of the target stimulus and the meaning extracted from this signal that is important in producing the interference effects. If the target

[1] For ease and clarity of presentation, we present the analyses of the RT data from Experiments 2–4 collapsed over block order. Neither the pattern of effects nor the conclusions drawn from these studies was at all influenced when block order was taken out as a between-subjects factor or when participants' data from each of the two block orders were analyzed separately.

requires a nondirectional decision, the processing of this stimulus and the parallel automatic processing of the task-irrelevant directional cue will result in the encoding of nonconflicting information, and, as a consequence, no interference effects will be observed. Therefore, interference should occur whenever participants are asked to make decisions based on direction, but nondirectional decisions should remain uninfluenced by to-be-ignored attentional cues. This hypothesis was examined in Experiment 3.

Experiment 3

In this experiment the pointing gestures used in Experiment 2 were replaced by "thumbs-up" and "thumbs-down" gestures. These are so-called "emblematic" or "symbolic" gestures that have a precise meaning understood by a particular group or culture and convey this meaning in the absence of any accompanying speech. Conventionally, throughout Europe the thumbs-up gesture means "good" or "okay," whereas the thumbs-down gesture means "bad" or "no good" (Morris, Collett, Marsh, & O'Shaughnessey, 1979). Thus, although these gestures are visually similar to the pointing gestures used in Experiments 1 and 2 (see Figure 1), they carry no intrinsic directional meaning. Again, participants were asked to respond to the gestures and head–gaze orientation in separate blocks of trials. However, instead of deciding whether the gesture was directed upward or downward, participants were asked to make one response if the gesture was thumbs-up and another if the gesture was thumbs-down. Because participants were not asked to make a decision based on the direction indicated by the gesture, responses were expected to be uninfluenced by the accompanying head–gaze orientation.

Method

Participants. Participants were 20 Open University students attending a residential summer school in Stirling. All had normal or corrected-to-normal vision.

Apparatus and materials. The visual stimuli consisted of digitized images of a male individual making a thumbs-up or thumbs-down gesture, in each case orienting his head upward or downward. Examples of these images are presented in Figure 2. These stimuli subtended approximately 9° of vertical visual angle and 7° of horizontal visual angle and were viewed by participants seated approximately 0.7 m from a 14-in. (35.56-cm) color monitor.

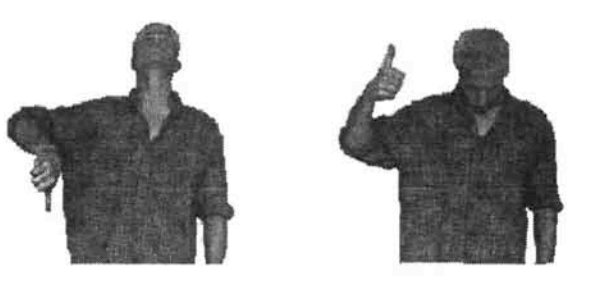

Figure 2. Reproductions of some of the incongruent stimuli used in Experiment 3.

Table 3
Mean Reaction Times (RTs; in Milliseconds) and Percentage of Errors for Responses to Head and Gesture Direction in Each Condition of Experiment 3

	Down targets		Up targets		Overall mean	
Congruity	RT	% of Errors	RT	% of Errors	RT	% of Errors
Responses to head						
Congruent	459	1.25	469	1.75	464	1.50
Incongruent	486	2.50	480	4.50	483	3.50
M	473	1.88	475	3.13	474	2.50
Responses to gesture						
Congruent	421	2.25	429	2.50	425	2.38
Incongruent	414	2.50	432	2.50	423	2.50
M	418	2.38	431	2.50	424	2.44

Design and procedure. The design and procedure were identical to those of Experiment 2, with one notable exception. In contrast to the gesture response condition of Experiment 2, in which participants based their responses on the direction indicated by the model's pointing gesture, participants in the corresponding condition of Experiment 3 were asked to decide whether the model was making a thumbs-up or a thumbs-down gesture. The instructions in the head response condition were identical to those of Experiment 2.

Results

The interparticipant means and percentage of errors for the eight conditions of Experiment 3 are shown in Table 3. When responding to the orientation of the head, participants were generally faster and more accurate in the congruent than incongruent condition. However, the RT difference was more marked when participants responded to the head oriented downward (27 ms) than upward (11 ms). In contrast, there was no consistent effect of congruity for responses to the gestures. Responses to thumbs-up gestures were just 3 ms faster in the congruent than incongruent condition, whereas RTs to thumbs-down gestures were actually 7 ms faster when the head was oriented in an incongruent than a congruent direction.

These observations were supported by a 2 (response dimension) × 2 (congruity) × 2 (decision) ANOVA conducted on participants' mean RTs. This analysis revealed main effects of congruity, $F(1, 19) = 5.98$, $p < .05$, response, $F(1, 19) = 29.86$, $p < .001$, and decision, $F(1, 19) = 4.55$, $p < .05$, and an interaction between response and congruity, $F(1, 19) = 14.00$, $p < .01$. However, these effects were qualified by a significant higher order interaction among the three variables, $F(1, 19) = 5.08$, $p < .05$. This interaction seemed to arise because of the large effect of congruity for responses to heads oriented downward compared with upward (27 vs. 11 ms) and the absence of any such congruity effect when participants responded to either type of gesture. Indeed, further analysis of this interaction indicated that the effect of congruity was significant for responses to heads oriented downward ($p < .001$), but only marginally significant for heads oriented upward ($p = .07$). There were no effects of congruity for responses to the gesture dimension ($p = .59$).

The overall mean error rate was only 2.47%. The correlation between mean RTs and mean error scores was .22, suggesting no evidence of a speed–accuracy trade-off. No further analysis was conducted on these data.

Discussion

The results of Experiments 1 and 2 establish that directional information obtained from the orientation of the head, the eyes, or both is processed automatically and is able to influence directional decisions to both spoken words and pointing gestures. In Experiment 3, however, these same head–gaze cues failed to influence nondirectional decisions to certain emblematic gestures that nevertheless contained visually similar information to the pointing gestures studied in Experiments 1 and 2. This finding supports the hypothesis that interference effects between signals will emerge only when a directional decision has to be made to any one of these cues.

However, when participants were asked to make a response to the head–gaze cues, the emblematic gestures did produce a reliable interference effect (marginally significant for "down" responses). Why should this be so, particularly when these gestures did not seem to be treated as directional signals when they formed the target stimulus dimension? One possibility is that when responding to the head–gaze dimension, the gestures would have appeared in the periphery of participants' vision and may therefore have been analyzed based only on low spatial frequency information, an analysis that might not have been sufficient to distinguish between a thumb and index finger. Thus, without the benefit of focal attention, the gestures could easily be processed as pointing signals and exert their effects on decisions to head–gaze direction accordingly.

On the other hand, it is possible that the default is to process these kinds of gestures as directional stimuli rather than assigning them their intended meanings. When these stimuli form the irrelevant dimension, it may be that it is their directional meanings that become active and hence exert an interfering effect on directional responses to head orientation. When these gestures form the target dimension, the demands of the task dictate that they are not processed directionally and so are uninfluenced by the irrelevant directional head cues. The asymmetrical nature of the effects in this experiment could therefore be attributable to some intrinsic directional meaning contained in these hand gestures, rather than their visual confusability with pointing gestures.[2]

Regardless of how the to-be-ignored gestures exerted their effects on responses to head orientation, the main finding of this experiment was that irrelevant head cues produced no interfering effect on nondirectional decisions to gestural stimuli. The elimination of this effect is notable given that the same cues produced reliable interference effects in Experiment 1, in which participants were responding to targets in an entirely separate modality, as well as in the intramodal task used in Experiment 2.

Experiment 4

Experiment 3 has established that nondirectional decisions to a social gestural signal are not influenced by directional information carried by head and gaze cues. In Experiment 4 we asked whether these same head–gaze cues would influence directional decisions to ostensibly nonsocial directional signals. It is possible, for example, that cues such as head and gaze orientation will produce interference effects only when participants have to make a decision based on cues that communicate something socially about direction (e.g., "Which way is he pointing?" "Which way is he telling me to look?"). However, it seems more likely that it is the directional nature of the decision that is important, rather than the social status of the directional cue. If participants are required to extract some directional meaning from any target and make a response contingent on this, then any other available directional cues may produce an interference effect as described earlier.

Thus, in Experiment 4, the pointing gestures used in Experiment 1 were replaced by arrows that were printed across the chest of the model. Again, participants were asked to make speeded keypress responses contingent on the direction of the arrow and the orientation of the head in separate blocks of trials. If the effects observed in Experiments 1 and 2 were caused by the social nature of the directional decision, then head cues would not be expected to interfere with responses to arrows. However, if it is the directional decision per se that is critical, then an interference effect should be observed.

Method

Participants. The participants were 20 Open University students attending the residential summer school. All had normal or corrected-to-normal vision, and none had participated in the previous experiments.

Apparatus and materials. The visual stimuli consisted of digitized images of the same male individual who served as a model for the stimuli used in Experiment 3. Images were obtained of him orienting his head either upward or downward as before but without making any accompanying hand gestures. These stimuli subtended approximately 8° of vertical visual angle and 6° of horizontal visual angle and again were viewed by participants

---

[2] The tendency to process the gesture cues directionally seems to be strong. Participants who first received the gesture response task knew that the gestures represented thumbs-up and thumbs-down, rather than pointing, and yet these stimuli still interfered with responses to head orientation in the second block of trials. This observation was confirmed by an additional ANOVA conducted on the RT data from Experiment 3, with block order as a between-subjects factor and response, congruity, and target direction as within-subjects factors. There was no main effect of block order ($p = .65$), nor were there any interactions involving this factor ($ps > .1$). In particular, the three-way interaction among block order, response, and congruity failed to reach significance ($p = .70$).

seated approximately 0.7 m from a 14-in. (35.56-cm) color monitor. For each head orientation, congruent and incongruent stimuli were created by pasting arrows, measuring approximately 3° × 2°, across the chest of the model, pointing either in the same direction or in the direction opposite to the orientation of his head. These images are shown in Figure 3.

Design and procedure. As in Experiment 1, the materials were tested in a 2 × 2 × 2 within-subjects design. The three variables were response (either to the arrow direction or to the orientation of the head), congruity (congruent or incongruent head–arrow pairings), and target direction (up or down). Participants were asked to respond to the direction of the arrow in one block of trials and to the orientation of the head in a second block. Half the participants responded to the arrow first and half to the head. The order of presentation within the blocks was completely randomized.

Results

Mean correct RTs and error rates in each condition of the experiment are presented in Table 4. When responding to arrows, participants were generally slower and less accurate in the incongruent than congruent condition. However, for head orientation, this RT difference was apparent only for heads oriented downward.

The RT data were entered into a 2 (response) × 2 (congruity) × 2 (target direction) ANOVA, which yielded a main effect of congruity, $F(1, 19) = 9.93, p < .01$, with RTs 13 ms slower in the incongruent than congruent condition. The analysis also yielded a marginally significant interaction between target decision and congruity, $F(1, 19) = 4.15, p = .06$, which seemed to arise because the congruity effect was larger for down targets (19 ms) than for up targets (7 ms). Although the data in Table 4 indicate that performance was equivalent for upward oriented heads paired with congruent and incongruent arrows, the three-way interaction term failed to reach significance ($p = .22$). No other main effects or interactions approached significance ($ps > .1$).

The overall mean error score in this experiment was only 1.19%. The correlation between the overall mean RTs and mean error rates in each condition was −.39. Thus, although the correlation failed to reach statistical significance, there was some evidence that participants were trading accuracy for speed. However, because the error rates in all conditions were so low, and generally mirrored the RT congruity effects, the small negative correlation was not deemed sufficient to compromise the interpretation of the RT data.

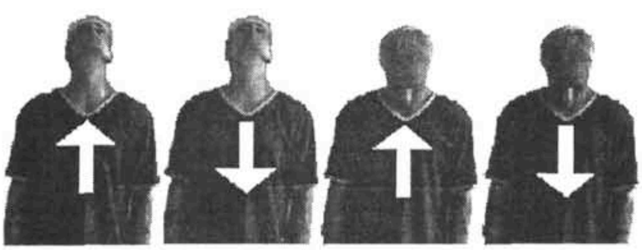

Figure 3. Reproductions of the digitized stimuli used in Experiment 4.

Table 4
Mean Reaction Times (RTs; in Milliseconds) and Percentage of Errors for Responses to Head and Arrow Direction in Each Condition of Experiment 4

	Down targets		Up targets		Overall mean	
Congruity	RT	% of Errors	RT	% of Errors	RT	% of Errors
Responses to head						
Congruent	425	2.25	438	0.25	432	1.25
Incongruent	445	1.50	438	2.00	442	1.75
M	435	1.88	438	1.13	437	1.50
Responses to arrow						
Congruent	438	0.75	439	0.75	439	0.75
Incongruent	457	1.00	452	1.00	455	1.00
M	448	0.88	446	0.88	447	0.88

Discussion

The most notable finding from this experiment is that to-be-ignored head–gaze cues produced significant interference effects on participants' responses to the arrow stimuli. It seems, then, that directional signals are able to exert effects on decisions to both social and nonsocial cues. Clearly, it is the directional meaning encoded from the target that is important in producing interference effects rather than the social nature of the decision or the social status of the target stimulus.

Although numerically only apparent for downward decisions, to-be-ignored arrows also exerted an interference effect on responses to head orientation. Langton et al. (1996) found that irrelevant arrows interfered with spoken words, so it is not all that surprising that the results of Experiment 4 provide some evidence that, like head and gesture cues, the direction of arrows seems to be processed automatically. However, it is unclear why this effect should occur only for downward decisions. One possibility is that with the head-down stimuli, the face and arrow are in closer proximity than in the head-up stimuli (see Figure 3). Directing attention toward the face may also bring the arrow into the attentional focus, which may in turn facilitate its processing.

General Discussion

In the studies reported in this article, we further examined the processing of pointing gestures of the type studied by Langton et al. (1996). Using a Stroop-type interference paradigm, Langton et al. found that to-be-ignored pointing gestures influenced the speed of responses to spoken directional words, suggesting that observers process these directional signals automatically. However, in addition to the outstretched hand, the directional information contained in the Langton et al. gestures was provided by the orientation of the head and the line of sight of the gesturer's gaze. Any or all of these signals could have received automatic processing and hence influenced responses to spoken words.

Experiment 1 in the present article replicated the Langton et al. (1996) finding. Again, to-be-ignored pointing gestures exerted an interference effect on responses to spoken directional words. However, this effect was found to be moderated by the direction in which the gesturer's head and gaze were oriented. Clearly, observers were automatically extracting directional information from the gesturer's head and eyes as well as from his hand. Moreover, the effects of head–gaze cues and gesture on spoken words were found to be additive, implicating a model in which pointing gesture and head cues are processed automatically and in parallel by separate systems. In support of this model, Experiment 2 demonstrated that head and gesture cues produced bidirectional interference effects in an intramodal version of the Stroop-type task. Participants' choice RTs to the direction of pointing gestures were slower when the gesturer's head and eyes were oriented in the direction opposite to the gesture compared with a condition in which gesture and head were oriented in the same direction. Reciprocally, responses to head angle were similarly influenced by to-be-ignored gestural information. When participants made nondirectional decisions to emblematic gestures in Experiment 3, interference effects were absent, suggesting that the locus of the interference effect between head and gesture cues is at a decision stage of processing following the independent evaluation of the cues. Finally, in Experiment 4, to-be-ignored head cues produced significant interference effects on responses to arrows, further suggesting that it is the nature of the directional decision that is important in producing the effects, rather than the social nature of the target stimulus.

It is clear from these findings and from experiments reported elsewhere (e.g., Driver et al., 1999; Friesen & Kingstone, 1999; Langton & Bruce, 1999; Langton et al., 1996) that various directional signals are processed automatically by observers. Why should this be so? One possibility is that these signals all carry information about the direction of another individual's *social attention.* In the remainder of this section, we discuss this hypothesis and speculate about the nature of the interaction or combination of information thought to produce the interference effects noted in Experiments 1–4.

Most organisms will orient their sensory receptors toward objects in the environment that are of immediate interest to them. Human and nonhuman primates, for example, will direct their eyes and heads toward individuals with whom they might interact, mate, or perhaps eat. To prepare for any of these encounters, it is critical that the potential interactant, sexual partner, or meal be able to detect when it has become the recipient of another's attention. Alternatively, rather than signaling its interest in the target, another's direction of attention might indicate the presence and the exact location of a third party or object, which again might constitute a threat or potential food source. By computing this individual's angle of gaze and then following this line of sight, the focus of its attention can be perceived and any necessary action taken. Along with alerting organisms to the presence of a predator or prey, this kind of "attention following" behavior is critical in human communication, such as in understanding otherwise ambiguous deictic expressions such as "here," "there," "that one," or "this one." Gaze following may also facilitate vocabulary acquisition by toddlers, as the referent of a new word can be specified by the direction in which the speaker is looking (Baldwin, 1991) or perhaps by pointing. In line with this, the results of several studies have shown that nonhuman primates (e.g., Emery, Lorincz, Perrett, Oram, & Baker, 1997), infants (e.g., Butterworth & Jarrett, 1991; Hood et al., 1998), and adults (Driver et al., 1999; Langton & Bruce, 1999) spontaneously redirect their gaze, their visual attention, or both in accord with another's gaze or head orientation.

Given their importance to survival and communication, it is perhaps unsurprising that people may have evolved mechanisms that automatically process directional cues such as head and gaze direction. Indeed, neuropsychological, neurophysiological, and behavioral evidence is emerging in support of the position that there is a functionally specific mechanism devoted to the task of detecting eyes and computing where in the environment eye gaze is directed (e.g., Campbell, Heywood, Cowey, Regard, & Landis, 1990; Heywood & Cowey, 1992; Perrett et al., 1988; Perrett et al., 1985). Moreover, Perrett and his colleagues suggested that something like a gaze detection mechanism forms only part of a system designed to process the direction of "social attention" (e.g., Perrett & Emery, 1994; Perrett, Hietanen, Oram, & Benson, 1992). Their work has indicated that individual cells in the superior temporal sulcus (STS) region of the macaque brain respond to *conjunctions* of eye, head, and body position. Thus, cells that respond preferentially to eyes directed downward show further preferences when the head is oriented downward and when the body adopts a quadrupedal posture. Accordingly, Perrett et al. (1992) suggested that one function of the STS region lies in the analysis of social attention direction. This mechanism, called the "direction of attention detector" by Perrett and Emery (1994), is considered to combine information from separate detectors analyzing body, head, and gaze direction to compute the whereabouts of another individual's focus of interest.

Pointing gestures also provide an important additional source of information about another's direction of attention. Indeed, they may provide a more accurate cue to the spatial location of a referent than either eye or head orientation (Butterworth, cited in Butterworth, 1995). Because of the value of these cues, it is possible that structures exist that process pointing gestures, along with gaze and head orientation as additional cues to the direction of social attention. Thus, it may well be that the automatic processing of head, gaze, and gestural information noted in the present experiments reflects the operation of some kind of direction of attention detector. However, further research is needed to establish whether participants are actually processing these directional signals as cues to social attention direction.

A possible argument against the suggestion that participants are processing social attention direction is the finding of Experiment 4 that arrows, like pointing gestures and head cues, also receive automatic processing. The argument is that arrows are conventional, nonsocial cues to direction and so would not be expected to be processed in the same way as social attention cues such as gaze, head orientation, or pointing gestures. The fact that arrows do produce an interfering effect on responses to head cues suggests that there is nothing special about these social signals.

However, the evidence that arrows received obligatory processing in Experiment 4 is equivocal, with numerical effects apparent when participants responded to heads directed downward but not upward. The qualitative difference between the effects of these to-be-ignored arrows and to-be-ignored social cues in Experiments 1 and 2 may suggest that arrows are processed rather differently from social signals. Alternatively, one might ask whether arrows are really conventional nonsocial symbols at all. Rather than acting as arbitrary directional symbols, arrows may have emerged as schematic representations of manual pointing gestures and might also be considered as social in the sense of communicating something about intention or meaning to another individual. In this respect, it is perhaps unsurprising that Experiment 4 revealed some evidence that arrows, like head, gaze, and gesture signals, are processed automatically.

Regardless of whether participants in Experiments 1–4 were processing social attention direction, the results of these experiments do indicate that the gestures used by Langton et al. (1996) actually comprised signals carried by the gesturer's head, gaze, and hand and that these cues were all processed automatically. The additional suggestion that these cues are initially processed in parallel and interact before the programming and execution of a response is similar to a model of information integration developed by Massaro and his colleagues (e.g., Massaro & Friedman, 1990). Their fuzzy logical model of perception involves three operations: evaluation, integration, and decision. Information from the two sources is evaluated independently and then integrated according to a multiplicative algorithm that ensures that the least ambiguous source of information carries the most weight in the decision process. Following integration, a decision is made on the basis of the relative goodness of match of the integrated stimulus information with the relevant prototype descriptions in memory. Massaro's group has applied this model to a variety of domains, including interactions between gesture and speech (Thompson & Massaro, 1986, 1994) and between emotional expressions in the face and voice (Massaro & Egan, 1996).

However, it may be that the interaction between attention signals observed in the present experiments is best accounted for by a *nonintegrative* interaction. In these experiments participants were asked to attend selectively to one relevant cue and to ignore information carried by the second irrelevant signal. On the whole, participants were able to do this successfully, suffering no loss in accuracy, only a reduction in speed, when the cues carried conflicting information. This suggests that participants' decisions were based on intact rather than integrated representations of information encoded from the target cue. Clearly, the precise mechanisms of the interaction require further investigation.

Another hypothesis currently under investigation is that the type of cross-modal interference effects noted in Experiment 2 here and by Langton et al. (1996) are mediated by the effect that certain social signals can exert on an observer's visuospatial attention. A number of research groups, including our own, have recently established that nonpredictive head–eye gaze cues (Langton & Bruce, 1999) and gaze cues from images of real faces (Driver et al., 1999) and schematic faces (Friesen & Kingstone, 1999) can trigger a reflexive, exogenous visual orienting response on behalf of an observer (see Spence & Driver, 1994, for a review of visual orienting). According to one particular theory of spatial attention, the "premotor theory" (e.g., Rizzolatti, Riggio, & Sheliga, 1994), to shift attention to a particular location entails the programming of an eye movement to that location, regardless of whether the eye movement is ever actually executed. This motor program contains directional features that becomes a spatial code when a manual response has to be selected. In this way, a stimulus containing a social attention cue would automatically generate a spatial code before initiating an orienting response. This spatial code could then somehow interfere with a spatial representation encoded from the spoken directional word at some stage in processing.[3] Whether this kind of mechanism could account for interference effects between social attention cues is not clear. However, the fact that head and gaze cues are able to trigger a reflexive orienting response on behalf of an observer, despite this observer's intentions and incentives to ignore these signals, provides further evidence that at least some social attention cues are analyzed rapidly and automatically by the information-processing system.

Finally, we emphasize that the visual stimuli used in the present experiments were all static images: poses of manual pointing, head, and eye gaze cues that would ordinarily be strongly dynamic events. In addition, gestures naturally occur in close synchronization with semantically corresponding spoken words (e.g., Condon, 1970; Levelt, Richardson, & La Heij, 1985). Clearly, researchers can only speculate that the effects we have observed here will transfer to naturally occurring moving stimuli, but given the attention-grabbing properties of such stimuli it is even possible that such effects will be augmented. Design constraints precluded the use of dynamic stimuli in the present studies, but a future challenge is to develop methodologies that could explore similar questions with more naturally occurring materials.

---

[3] This argument is similar to and is motivated by an account of the Simon effect (e.g., Simon & Rudell, 1967) given by Nicoletti, Umiltà, and their colleagues (e.g., Nicoletti & Umiltà, 1994; Rubichi, Nicoletti, Iani, & Umiltà, 1997). Briefly, the Simon effect occurs when the location of a target stimulus interferes with the response to that stimulus. Typically, participants are asked to make a left–right keypress response contingent on the identity of a stimulus (e.g., color of light, direction of arrow, the words *left* or *right*, etc.) presented randomly to the left or right of some central point. Responses are faster when the location of the stimulus matches the location of the response than when it does not.

References

Argyle, M. (1988). *Bodily communication* (2nd ed.). London: Methuen.

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

Butterworth, G. (1995). Origins of mind in perception and action. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 29–40). Hillsdale, NJ: Erlbaum.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9,* 55–72.

Campbell, R., Heywood, C. A., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia, 28,* 1123–1142.

Condon, W. S. (1970). Method of micro-analysis of sound films of behavior. *Behavioral Research Methods and Instrumentation, 2,* 51–54.

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61–83). Hillsdale, NJ: Erlbaum.

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.

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

Feyereisen, P. (1991). Brain pathology, lateralization, and nonverbal behavior. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 31–70). Cambridge, England: Cambridge University Press.

Friesen, C. K., & Kingstone, A. (1999). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review, 5,* 490–495.

Heywood, C. A., & Cowey, A. (1992). The role of the "face cell" area in the discrimination and recognition of faces by monkeys. *Philosophical Transactions of the Royal Society of London Series B, 335,* 31–38.

Hood, B. M., Willen, J. D., & Driver, J. (1998). An eye direction detector triggers shifts of visual attention in human infants. *Psychological Science, 9,* 131–134.

Kendon, A. (1986). Some reasons for studying gesture. *Semiotica, 62,* 3–28.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction, 27,* 175–200.

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

Langton, S. R. H. (in press). The mutual influence of gaze and head orientation in the analysis of social attention direction. *Quarterly Journal of Experimental Psychology.*

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., O'Malley, C., & Bruce, V. (1996). Actions speak louder than words: Symmetrical cross-modal interference effects in the processing of verbal and gestural information. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1357–1375.

Levelt, W. J. M., Richardson, G., & La Heij, W. (1985). Pointing and voicing in deictic expressions. *Journal of Memory and Language, 24,* 133–164.

Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the face and voice. *Psychonomic Bulletin and Review, 3,* 215–221.

Massaro, D. W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review, 97,* 225–252.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746–748.

McNeill, D. (1985). So you think gestures are non-verbal? *Psychological Review, 92,* 350–371.

McNeill, D. (1987). So you *do* think gestures are non-verbal! Reply to Feyereisen (1987). *Psychological Review, 94,* 499–504.

McNeill, D. (1989). A straight path to where? Reply to Butterworth and Hadar. *Psychological Review, 96,* 175–179.

McNeill, D., Cassell, J., & McCullough, K. (1994). Communicative effects of speech-mismatched gestures. *Research on Language and Social Interaction, 27,* 223–237.

Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode and practice. *Memory & Cognition, 21,* 627–645.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 615–622.

Morris, D., Collett, P., Marsh, P., & O'Shaughnessey, M. (1979). *Gestures, their origins and distribution.* London: Cape.

Nicoletti, R., & Umiltà, C. (1994). Attention shifts produce spatial stimulus codes. *Psychological Research, 56,* 144–150.

Perrett, D. I., & Emery, N. J. (1994). Understanding the intentions of others from visual signals: Neurophysiological evidence. *Cahiers de Psychologie Cognitive, 13,* 683–694.

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

Perrett, D. I., Mistlin, A. J., Chitty, A. J., Harries, M., Newcombe, F., & de Haan, E. (1988). Neuronal mechanisms of face perception and their pathology. In C. Kennard & F. Clifford Rose (Eds.), *Physiological aspects of clinical neuro-ophthalmology* (pp. 137–154). London: Chapman & Hall.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London, B223,* 293–317.

Rimé, B., & Schiaratura, L. (1991). Gesture and speech. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 239–281). Cambridge, England: Cambridge University Press.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 231–265). Hillsdale, NJ: Erlbaum.

Rubichi, S., Nicoletti, R., Iani, C., & Umiltà, C. (1997). The Simon effect occurs relative to the direction of an attention shift. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1353–1364.

Simon, J. R., & Rudell, A. P. (1967). The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51,* 300–304.

Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 555–574.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Thompson, L. A., & Massaro, D. W. (1986). Evaluation and integration of speech and pointing gestures during referential understanding. *Journal of Experimental Child Psychology, 42,* 144–168.

Thompson, L. A., & Massaro, D. W. (1994). Children's integration of speech and gestures in comprehension. *Journal of Experimental Child Psychology, 57,* 327–354.

Received January 28, 1998
Revision received February 15, 1999
Accepted May 17, 1999 ■