Final author draft only – the final article published in *Memory* may differ slightly from this manuscript version.

Look who's talking! Facial appearance can bias source monitoring

Robert A. Nash, Olwen M. Bryer, & Friederike Schlaghecken

University of Warwick, Coventry, UK

Word count: 3401 (main text)

Corresponding author:

Robert A. Nash
Department of Psychology
University of Warwick
Coventry, CV4 7AL
United Kingdom
Direct: +44 24 7652 8386   Fax: +44 24 7652 4225
Email: R.Nash@warwick.ac.uk

Abstract

When we see a stranger's face, we quickly form impressions of his or her personality, and expectations of how the stranger might behave. Might these intuitive character judgments bias source monitoring? Subjects read headlines 'reported' by a trustworthy- and an untrustworthy-looking reporter. Subsequently, subjects recalled which reporter provided each headline. Source memory for likely-sounding headlines was most accurate when a trustworthy-looking reporter had provided the headlines. Conversely, source memory for unlikely-sounding headlines was most accurate when an untrustworthy-looking reporter had provided the headlines. This bias appeared to be driven by the use of decision criteria during retrieval rather than differences in memory encoding. Nevertheless, the bias was apparently unrelated to variations in subjective confidence. These results show for the first time that intuitive, stereotyped judgments of others' appearance can bias memory attributions analogously to the biases that occur when people receive explicit information to distinguish sources. We suggest possible real-life consequences of these stereotype-driven source monitoring biases.

Look who's talking! Facial appearance can bias source monitoring

Being able to accurately identify how we acquired a particular piece of information is important in many aspects of everyday life (Davis & Friedman, 2007; Sherman & Bessenoff, 1999). One reason is that when we know from where information originated, we can assess its reliability (Gordon, Franklin, & Beck, 2005): for instance, information learnt from a trusted friend is probably more likely to be accurate than is information learnt from a reputed liar. Unfortunately, though, our ability to match our memories to their sources is fallible. Because memories are not stored with convenient 'labels' indicating their origin, when we remember we engage in an attribution process known as *source monitoring*. That is, we infer the origin of information—correctly or incorrectly—by relying on additional source cues (Johnson, Hashtroudi, & Lindsay, 1993; Lindsay, 2008).

Most source monitoring research has focused on a single variety of source cue—the phenomenological qualities of memories, such as their perceptual vividness and cohesiveness. However, researchers have more recently explored how people use a second variety of cue—their background knowledge about different sources' characteristics (Bayen, Nakamura, Dupuis, & Yang, 2000). For example, Fragale and Heath's (2004, Study 2) subjects read a story describing evidence pertaining to a murder investigation. Each piece of evidence incriminated one of two suspects, and was described as having been reported in either a newspaper or a gossip sheet. After subjects learnt which suspect was factually guilty, they recalled which evidence originated from each source. Subjects showed a bias toward attributing 'correct' evidence—which incriminated the guilty suspect—to the newspaper (the more reliable source), and 'incorrect' evidence to the gossip sheet (the less reliable source).

In a similar way, our knowledge about social categories and characteristics, such as a person's gender, sexual orientation, or occupation, can also bias source attributions (Cook, Marsh, & Hicks, 2003; Marsh, Cook, & Hicks, 2006; Sherman & Bessenoff, 1999). For

instance, in a study by Mather, Johnson and De Leonardis (1999), subjects saw video-clips of two speakers making political statements and subsequently received information that characterized each speaker as either Democrat or Republican. When subjects tried to recall who made particular statements, they were biased toward attributing statements matching a 'Democrat schema' to the Democrat speaker, and statements matching a 'Republican schema' to the Republican speaker.

However, in each of these studies, source information was given to subjects explicitly; that is, sources were labelled as either 'newspaper' or 'gossip sheet,' as either 'Democrat' or 'Republican,' and so forth. In the present study, we ask whether people use attribution biases even when they receive no explicit information about sources' characteristics. Specifically, we ask whether people's intuitive character judgements—based on nothing but sources' facial appearance—can guide their source monitoring.

On the one hand, facial appearance could in principle be sufficient to elicit such an attribution bias. It is well known that people form strong and lasting impressions of others based solely on facial characteristics (for a recent review see Todorov, Said, Engell, & Oosterhof, 2008). When we see a stranger's face, we automatically and almost instantaneously categorize them, not only by considering relatively unambiguous attributes such as sex and race (Stangor, Lynch, Duan, & Glass, 1992; Taylor, Fiske, Etcoff, & Ruderman, 1978), but also by inferring personality characteristics such as competence, likeability or trustworthiness (Todorov, Mandisodza, Goren & Hall, 2005; Willis & Todorov, 2006). These character judgements—though not necessarily accurate—guide people's expectations and interpretations of others' behaviour in many aspects of everyday life (Bull & Green, 1980; Bull & Hawkes, 1982; Hochberg & Galper, 1974). In short, the ubiquity and automaticity of facial stereotyping might lead one to expect that facial appearance would be sufficient to bias source monitoring.

On the other hand, there are several reasons to believe that facial characteristics would not elicit attribution biases. First, attribution bias effects are mainly observed when the characteristics of different sources are highly salient (Davis & Friedman, 2007), whereas one might argue that with human faces, personality characteristics are inferred from rather subtle cues. Second, even with explicitly labelled, clearly distinct sources, pronounced effects are mostly found under poor memory-encoding or -retrieval conditions. For instance, biases are strongest among older adults (Mather et al., 1999), or when subjects experience high cognitive load during encoding or retrieval (Ehrenberg & Klauer, 2005; Sherman & Bessenoff, 1999). Third, it has been shown that source attribution biases are most pronounced when information about the sources is available at retrieval but not during encoding (Cook et al., 2003; Hicks & Cockman, 2003). Clearly, this condition is not fulfilled if the source is represented by (a photograph of) a human face: Because facial stereotypes are formed almost instantaneously upon exposure (Willis & Todorov, 2006), any characteristics inferred from faces should be inferred during encoding, thus potentially minimizing any bias effects.

The following experiment investigated this issue with respect to one specific facial cue—the perceived trustworthiness of sources' faces. Subjects read fictional news headlines reported by a trustworthy-looking and an untrustworthy-looking reporter. After a short delay, they tried to recall which reporter provided each headline. We were primarily interested in whether people would tend to attribute likely headlines (i.e., whose outcome seems probable) to the more trustworthy-looking reporter, and unlikely headlines to the untrustworthy-looking reporter. Hereafter, we refer to this prediction as an *inferential bias*.

However, people are influenced not only by what they rationally infer must be true, but also by what they would like to be true. Such 'wishful thinking' has been shown to produce source monitoring errors similar to those described above (Barber, Gordon, & Franklin, 2009; Gordon et al., 2005). Thus a secondary interest of the present study was to examine whether

facial trustworthiness would cue a *wishful thinking bias,* whereby subjects would tend to attribute desirable headlines (i.e., whose outcome is desirable) to the more trustworthy-looking reporter, and undesirable headlines to the untrustworthy-looking reporter. If a wishful thinking bias were to occur, it might at least partly counter any inferential bias when likelihood and desirability are at odds (i.e., for likely-undesirable and unlikely-desirable headlines). Thus we also predicted a three-way interaction upon source accuracy between reporter's appearance, headline likelihood, and headline desirability.

Method

Subjects & Design

Forty undergraduates (53% female, $M_{Age} = 20.85$, $SD = 1.86$, Range = 19-30) participated without payment. The study used a 2 (Reporter's appearance: trustworthy vs. untrustworthy) x 2 (Headline likelihood: likely vs. unlikely) x 2 (Headline desirability: desirable vs. undesirable) repeated-measures design.

Materials

Face stimuli

From the Essex Face Database (Spacek, n.d.) we selected 40 facial photographs of young Caucasian males without glasses, piercings or facial hair, in neutral expression and frontal view. Forty non-participating volunteers rated the trustworthiness of each face (1 = very untrustworthy; 7 = very trustworthy), amongst other traits (see General Discussion). We selected the four photographs with the highest mean trustworthiness ratings (combined $M = 4.78$), and the four with the lowest mean ratings (combined $M = 2.79$) to use in the main study. We paired these eight photographs such that the members of each pair differed significantly in terms of trustworthiness (all $ps < .001$). Subjects in the main study were randomly assigned to see one of the four photograph-pairs.

Headlines

We created forty 9- to 13-word fictional news headlines; ten for each of the four cells in our 2 (Likelihood) x 2 (Desirability) manipulation, and based on our own judgments of likelihood and desirability. As a manipulation check, the 40 non-participating volunteers rated the likelihood (1 = highly unlikely; 7 = highly likely) and desirability (1 = very undesirable; 7 = very desirable) of each headline-event. From each cell, we removed two headlines whose mean ratings were close to, or on the unexpected side of, the midpoint of either scale. The remaining likely headlines were judged as more likely than unlikely headlines ($M = 4.64$ vs. $M = 2.50$, $p < .001$), and desirable headlines were judged as more desirable than undesirable headlines ($M = 5.69$ vs. $M = 1.58$, $p < .001$). These 32 headlines were used in the main study, and examples are provided in Table 1. The full set of headlines is available from the authors.

[INSERT TABLE 1 ABOUT HERE]

Procedure

We told subjects that they would read several headlines provided by two junior newspaper reporters. Subjects previewed the photographs of the two reporters they had been assigned: one trustworthy-looking, one untrustworthy-looking. Next, subjects were instructed to imagine they were the newspaper's editor, and were told that they would later need to recall which reporter provided each headline. They saw a series of 32 slides on a computer screen, each comprising a headline on the right of the screen in 32-point font, and a large photograph of a reporter on the left. Each slide appeared for 10 s followed by a 1-s blank interval; the entire slide-sequence thus lasted just under 6 min. All subjects saw the same headlines in a randomized order. Subjects saw half of each headline-type (i.e., unlikely-desirable; likely-undesirable, etc.) paired with a trustworthy-looking face, and half of each type paired with an untrustworthy-looking face; the headline/face-type pairings were counterbalanced across subjects.

Next, subjects completed an anagram-solving filler task for 10 min. Finally they received a questionnaire, on which all 32 headlines were listed in a single randomized order, and the appropriate two reporters were pictured at the top of each page. Subjects recorded which reporter they believed provided each headline, by ticking a box underneath the appropriate photograph and rating their confidence in each response (1 = Guess; 5 = Certain).

Results and Discussion

In determining which reporter provided each headline, our subjects performed at 74.2% accuracy on average (Range = 50-100%). To avoid ceiling effects in our analyses, we removed 9 subjects from analyses whose accuracy exceeded 90%; removing these data did not change the overall pattern of findings.

Hereafter, we use the terms *congruent* and *incongruent* when referring to our experimental predictions, using a valence-based definition of congruence. That is, trustworthy faces are congruent with likely and desirable headlines, and untrustworthy faces are congruent with unlikely and undesirable headlines.

Source Accuracy

Data were analysed using a 2 (Reporter's appearance) x 2 (Headline likelihood) x 2 (Headline desirability) repeated-measures ANOVA on subjects' source accuracy scores. We found no main effects of headline likelihood, $F(1,30) = 0.27$, $p = .87$, $\eta_p^2 < .01$, or headline desirability, $F(1,30) = 0.05$, $p = .83$, $\eta_p^2 < .001$. However, the main effect of reporter's appearance approached significance, $F(1,30) = 3.48$, $p = .07$, $\eta_p^2 = .10$, showing that source accuracy was somewhat greater for trustworthy-looking than for untrustworthy-looking reporters' headlines.

Our primary question of interest was whether subjects used reporters' facial appearance to cue an inferential bias. A significant two-way interaction between reporter's appearance and headline likelihood was indeed observed, $F(1,30) = 9.24$, $p < .01$, $\eta_p^2 = .24$: source

accuracy was greater when the headline's likelihood was congruent rather than incongruent with the reporters' facial trustworthiness. In contrast, the wishful thinking bias—that is, the interaction between reporter's appearance and headline desirability—was not significant, although subjects' accuracy scores tended in the predicted direction, $F(1,30) = 2.52$, $p = .12$, $\eta_p^2 = .08$: the source of headlines was recalled somewhat more accurately when their desirability was congruent rather than incongruent with the reporters' facial trustworthiness. The interaction between headline likelihood and desirability was not significant, $F(1,30) = 0.13$, $p = .73$, $\eta_p^2 < .01$.

However, the significant interactions were qualified by a marginally significant three-way interaction, $F(1,30) = 3.03$, $p = .09$, $\eta_p^2 = .09$. As Figure 1 illustrates, source attributions for likely-desirable headlines were strongly biased by reporter's appearance ($d = 1.25$), whereas there was only a weak bias for all other headline types (largest $d = 0.25$). We return to consider this interaction in the General Discussion.

[INSERT FIGURE 1 ABOUT HERE]

Signal Detection Analysis

One possible mechanism underlying the observed inferential bias is that congruence between a reporter's appearance and headline likelihood facilitated encoding of source information in memory. An alternative, but not mutually exclusive, possibility is that subjects used a decision strategy during retrieval to help them to determine the source of headlines. To differentiate between these two mechanisms, we conducted signal detection analyses on source accuracy scores. The analyses' outcomes were highly similar regardless of whether we focused on accuracy for trustworthy- or for untrustworthy-looking reporters, thus here we arbitrarily report accuracy measures for the untrustworthy-looking reporters.

We predicted—following Gordon et al. (2005)—that if subjects' responses reflected differential encoding of congruent versus incongruent information, then there should be

differences in sensitivity (d') between likely and unlikely scenarios. If subjects' responses reflected a decision strategy at retrieval, then there should be differences in the criterion parameter (C). Following Gordon et al., we adjusted the proportions of hits and false alarms by converting X/N (where X = number of correct responses, and N = number of responses) to $(X + 0.5)/(N + 1)$. This adjustment ensured that cells containing accuracy scores of 0% or 100% did not produce infinite values for d' or C.

Comparing the mean signal detection parameters for likely and unlikely headlines, we found no significant differences in d' (M = .89 vs. M = .94), $t(30)$ = 0.30, p = .76, yet there were significant differences in C (M = .62 vs. M = .38), $t(30)$ = 2.16, p = .04. In other words, subjects adopted a stricter criterion for attributing likely headlines than for attributing unlikely headlines to the untrustworthy-looking reporter. In line with the findings of other source monitoring studies (e.g., Cook et al., 2003; Gordon et al., 2005), these analyses suggest that the inferential bias was driven by decision-processes during retrieval, and not by differences in encoding.

Confidence Ratings

We finally asked whether subjects' proneness to using this decision-based inferential bias was affected by their subjective confidence. It seems plausible that an inferential bias might only occur among specific subjects—for example, those who believe they have a poor memory for source information (general confidence). Alternatively, all subjects might use an inferential bias, but only when they are uncertain of the source of specific headlines (item-specific confidence).

First, with respect to general confidence, we calculated the mean confidence rating assigned by each subject across all 32 headlines. We then calculated the extent to which each subject demonstrated an inferential bias, by subtracting their mean source accuracy for likelihood-incongruent headlines from that for likelihood-congruent headlines. This measure

of bias did not correlate with subjects' mean confidence ratings, $r(29) = .01$, $p = .95$; that is, there was no relationship between subjects' general confidence and their tendency to use an inferential bias.

Second, we found scant evidence that subjects' item-specific confidence drove the inferential bias. A 2 (Reporter's appearance) x 2 (Headline likelihood) x 2 (Headline desirability) repeated-measures ANOVA on subjects' individual confidence ratings revealed no significant main effect of, nor interactions involving, reporter's appearance, and no main effects of headline likelihood or desirability, all ps > .22, largest $\eta_p^2 = .05$. There was, however, a marginally-significant interaction between headline likelihood and desirability, $F(1,30) = 4.09$, $p = .052$, $\eta_p^2 = .12$. Surprisingly, headlines with congruent likelihood and desirability (i.e., likely-desirable or unlikely-undesirable) were attributed with less confidence than those with incongruent features. Indeed, the headlines that attracted the lowest confidence ratings were likely-desirable headlines—which were the ones that were most noticeably biased by reporters' appearance (see Figure 1).

Although this finding fits with the hypothesis that subjects used decision-biases when they had low confidence in the source of a specific headline, we are reluctant to conclude that it can explain our results. First, the magnitudes of the differences were small: just 0.19 scale-points separated the headline-types with lowest and highest confidence ratings (Likely-desirable headlines, $M = 2.89$, $SD = 0.91$; Likely-undesirable headlines, $M = 3.08$, $SD = 0.93$). Second, aside from the lowest confidence ratings being attributed to the headline-type showing the greatest level of bias, there was no clear further relationship between these ratings and the level of bias observed. Third, and perhaps most importantly, when subjects made errors they were no less confident in their likelihood-congruent choices ($M = 2.56$) than their likelihood-incongruent choices ($M = 2.50$), $p = .65$. In sum, we conclude that differences

in subjects' overall and item-specific confidence cannot adequately explain the observed inferential bias.

General Discussion

The present study shows that when people recall the source of verbal information, they use the facial characteristics of possible sources as a decision-cue. To our knowledge, this is the first empirical demonstration that facial appearance can inform and systematically bias source monitoring. Following studies which show that facial judgments are made automatically (Willis & Todorov, 2006), it is noteworthy that this bias occurred without subjects being instructed to form impressions of the reporters.

Recall that our primary interest was in whether subjects would use reporters' appearance to cue an inferential bias. Indeed, when our subjects saw likely-sounding headlines, they were more accurate at subsequently recalling who reported them if the actual source was trustworthy- rather than untrustworthy-looking. At the bias's most prominent, subjects were almost 25% more likely to correctly identify the source of likely-desirable headlines when the reporter was trustworthy- than when he was untrustworthy-looking. Of course, in a human face—in contrast to artificially-constructed source information— 'trustworthiness' is not an isolated feature. People's judgements of numerous valence-based personality traits are often intercorrelated (Oosterhof & Todorov, 2008), and indeed, our non-participating volunteers' trustworthiness ratings were strongly related to their ratings of attractiveness ($r = .59$) and intelligence ($r = .77$). One might speculate that the results observed in the present study reflect a general bias towards linking positive (likely, desirable) information with a positive-looking (trustworthy, attractive, intelligent) source, and negative information with a negative-looking source. Alternatively, one might argue that these results reflect a specific effect of perceived trustworthiness, independently of other correlated traits. Further work is needed to tease these alternatives apart.

Overall, our findings fit with those of analogous studies that used explicit schematic information to differentiate sources (e.g., Fragale & Heath, 2004; Sherman & Bessenoff, 1999), thus providing an initial demonstration that even implicit stereotypes based on intuitive judgments can bias source monitoring. Accordingly, we predict that facial appearance would cause even larger biases under poorer encoding and/or retrieval conditions. It is of particular interest that the inferential bias appeared to have no relationship with subjective confidence. In contrast, Hicks and Cockman (2003) found that decision-biases were most prominently associated with high confidence; yet our findings do at least concur insofar as they show that source biases are not necessarily—as one might predict—a product of low confidence.

One thus far unexplained outcome of our study was that although subjects demonstrated a large attribution bias for likely-desirable headlines, the pattern was much less pronounced for other headline-types. Why would this be? As we predicted, the inferential bias could have been counteracted by an opposing wishful thinking bias in the cases of likely-undesirable and unlikely-desirable headlines. Yet this reasoning does not explain why we did not see larger biases in subjects' attributions of undesirable-unlikely headlines. It is known that source monitoring of negatively-valenced information is often superior to that of positively-valenced information (Bell & Buchner, in press; Mather, Shafir, & Johnson, 2000), but even this phenomenon cannot account for our results, as subjects' overall source accuracy was equivalent across all headline-types. This is not the first study to find larger attribution biases for positively- than for negatively-valenced information with equivalent overall source accuracy (e.g., Benney & Henkel, 2006); however, the mechanisms underlying this asymmetry and the circumstances under which it might occur are yet to be discovered.

Beyond their theoretical implications for source monitoring and impression formation, the present findings also have practical implications. Davis and Friedman (2007), for

example, propose that remembering 'who said what' might often have repercussions, such as

when a witness assists police by recalling who they heard make an incriminating statement.

Guided by our results, it seems possible that in such a scenario, a witness might falsely

attribute the statement to somebody with an untrustworthy, unattractive, or 'criminal-looking'

face. Indeed, future research should explore the extent to which different facial

characteristics, including criminality, influence source monitoring. Moreover, given that

stereotype-driven thinking can induce distortions of memory (e.g., Kleider, Goldinger, &

Knuycky, 2008; Ottati, Claypool, & Gingrich, 2005; Snyder & Uranowitz, 1978), it seems

possible that facial stereotypes could lead people to remember others doing things that never

truly occurred. This is an important issue for future research to consider.

References

Barber, S. J., Gordon, R., & Franklin, N. (2009). Self-relevance and wishful thinking:

Facilitation and distortion in source monitoring. *Memory & Cognition, 37,* 434-446.

Bayen, U. J., Nakamura, G. V., Dupuis, S. E., & Yang, C.-L. (2000). The use of schematic

information about sources in source monitoring. *Memory & Cognition, 28*, 480-500.

Bell, R., & Buchner, A. (in press). Valence modulates source memory for faces. *Memory &

Cognition.*

Benney, K. S., & Henkel, L. A. (2006). The role of free choice in memory for past decisions.

Memory, 14, 1001-1011.

Bull, R. H. C., & Green, J. (1980). The relationship between physical appearance and

criminality. *Medicine, Science, and the Law, 20*, 79-83.

Bull, R. H. C., & Hawkes, C. (1982). Judging politicians by their faces. *Political Studies, 30*,

95-101.

Cook, G. I., Marsh, R. L., & Hicks, J. L. (2003). Halo and devil effects demonstrate valence-

based influences on source-monitoring decisions. *Consciousness & Cognition, 12*, 257-

278.

Davis, D., & Friedman, R. D. (2007). Memory for conversation: The orphan child of witness

memory researchers. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.),

Handbook of eyewitness psychology (Vol. 2): Memory for events (pp. 3-57). Mahwah, NJ:

Lawrence Erlbaum Associates.

Ehrenberg, K., & Klauer, K. C. (2005). Flexible use of source information: Processing

components of the inconsistency effect in person memory. *Journal of Experimental

Social Psychology, 41*, 369-387.

Fragale, A. R., & Heath, C. (2004). Evolving informational credentials: The (mis)attribution of believable facts to credible sources. *Personality and Social Psychology Bulletin, 30*, 225-236.

Gordon, R., Franklin, N., & Beck, J. (2005). Wishful thinking and source monitoring. *Memory & Cognition, 33*, 418-429.

Hicks, J. L., & Cockman, D. W. (2003). The effect of general knowledge on source memory and decision processes. *Journal of Memory and Language, 48*, 489-501.

Hochberg, J., & Galper, R. E. (1974). Attribution of intention as a function of physiognomy. *Memory & Cognition, 2*, 39-42.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28.

Kleider, H. M., Goldinger, S. D., & Knuycky, L. (2008). Stereotypes influence false memories for imagined events. *Memory, 16*, 97-114.

Marsh, R. L., Cook, G. I., & Hicks, J. L. (2006). Gender and orientation stereotypes bias source-monitoring attributions. *Memory, 14*, 148-160.

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology, 16*, 437-458.

Mather, M., Shafir, E., & Johnson, M. K. (2000). Misremembrance of options past: Source monitoring and choice. *Psychological Science, 11*, 132-138.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*, 11087-11092.

Ottati, V., Claypool, H. M., & Gingrich, B. (2005). Effects of a group stereotype on memory for behaviors performed by a group member. *European Journal of Social Psychology, 35*, 797-808.

Sherman, J. W., & Bessenoff, G. R. (1999). Stereotypes as source-monitoring cues: On the interaction between episodic and semantic memory. *Psychological Science, 10*, 106-110.

Snyder, M., & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology, 36*, 941-950.

Spacek, L. (n.d.). University of Essex Face Database. Available at http://cswww.essex.ac.uk/mv/allfaces/index.html.

Stangor, C. Lynch, L., Duan, C., & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62*, 207-218.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*, 778-793.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*, 1623-1626.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences, 12*, 455-460.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592-598.

Author note

Table 1. Examples of headlines used in the main study.

	Likely	Unlikely
Desirable	"London Olympics investment will bring improvements in public transport in the capital"	"80% of kids give up TV for exercise after successful educational campaign"
Undesirable	"Researchers admit that AIDS cure may still be many years away"	"20% of UK prisoners to be released to make room in packed prisons"

Table 2. Main effects and interactions of experimental variables on subjects' confidence ratings.

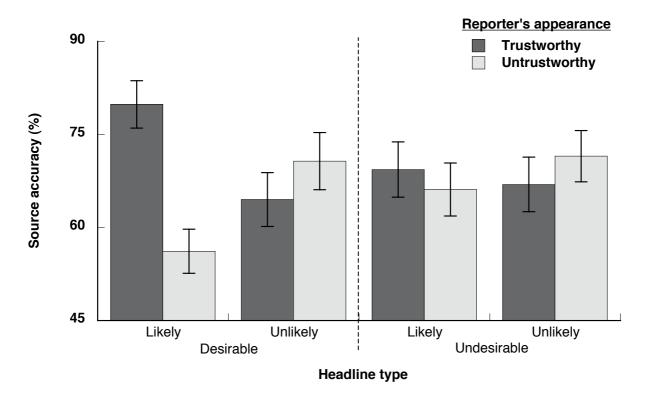| Variable(s)	F-statistic	p-value	Effect size (η^2_p)
Reporter's appearance	.48	.49	.02
Headline likelihood	.23	.64	.01
Headline desirability	.97	.33	.03
Reporter's appearance x headline likelihood	1.54	.22	.05
Reporter's appearance x headline desirability	.40	.53	.01
Headline likelihood x headline desirability	4.09	.05	.12
Reporter's appearance x headline likelihood x headline desirability	1.54	.22	.05

Figure 1.

Figure captions

Figure 1. Source accuracy as a function of face trustworthiness, headline likelihood, and headline desirability, ±1*SE*.