

Smiles behind a mask are detectable and affect judgments of attractiveness, trustworthiness, and competence[☆]

Astrid Hopfensitz^{a,*}, César Mantilla^b

^a EM-Lyon Business School, France

^b Economics Department, Universidad del Rosario, Colombia

ARTICLE INFO

Keywords:

Smiling
Non-verbal communication
COVID
Face masks
Halo effect

ABSTRACT

Smiling is a popular and powerful facial signal used to influence how we are judged and evaluated by others. The recent COVID pandemic made the use of face masks common around the world. Since face masks, when properly worn, cover the lower half of the face, a common concern is that they inhibit our ability to signal to others through facial expressions like smiles. In this paper, we show through three subsequent studies that smiling faces are easily distinguished from neutral faces even if the person is wearing a face mask (Study 1, $N = 1814$). We further show that smiling behind a face mask significantly influences ratings regarding attractiveness, trustworthiness, and competence (Study 2, $N = 250$). We finally show that individuals with about 18 months of experience with face masks are well aware that smiling behind face masks will influence ratings regarding attractiveness and trustworthiness by others (Study 3, $N = 94$). Together, our studies provide evidence that face masks should not be seen as a threat that inhibits simple non-verbal communication through smiles.

1. Introduction

The recent COVID pandemic has crucially altered how we interact with others. Government regulations worldwide and individual preferences to protect from the virus have led to an unprecedented increase in the use of hygienic face masks in public (Egan et al., 2021; Feng et al., 2020; Tso & Cowling, 2020). Proper use of face masks requires these to simultaneously cover the bearer's mouth and the nose, crucially altering what we see when we interact with others. While the use of face masks might reduce once the pandemic is controlled in larger parts of the world, it seems likely that they will remain much more present in shared spaces. It was the case in Asian countries, since the pneumonic plague outbreak in Manchuria, China, in 1910–11, and the 1918–19 flu pandemic in Japan (Burgess & Horii, 2012).

Meanwhile, individuals and professionals might wonder whether non-verbal communication and our tendency to form judgments of others based on their faces will be influenced by face masks (Marler & Ditton, 2021; Nobrega et al., 2020). This paper investigates to which degree observers can detect non-verbal signals from individuals wearing a face mask, notably smiles, and whether these signals, even when wearing a face mask, influence evaluations with respect to attractiveness, trustworthiness, and competence. Finally, the paper studies to which degree individuals are aware of the effect that face masks have on the perceptions regarding these traits.

Humans have a unique ability to use intricate social signals, and many of these signals arise from seeing others, especially their

[☆] Data, scripts, and supplementary online material can be accessed through <https://osf.io/hj3p5/>.

* Corresponding author.

E-mail addresses: hopfensitz@em-lyon.com (A. Hopfensitz), cesar.mantilla@urosario.edu.co (C. Mantilla).

faces. Human infants already focus on faces longer than other objects (Frank et al., 2009), and adults detect facial emotional expressions very fast (Sato & Yoshikawa, 2010). Facial expressions are employed to infer intentions, attention, and character judgments about others. While character judgments based on facial expressions are mostly wrong, they are predictable. Extensive literature in psychology, political science, and human resources has evidenced that specific facial expressions in others can influence our judgments. These judgments regarding others' attitudes, intentions, or expected behaviors ultimately affect behavior regarding cooperation (Orn et al., 2001; Stirrat & Perrett, 2010; Van Leeuwen et al., 2017), mate choice (Rhodes, 2006), hiring (Filkuková & Jørgensen, 2021), and voting (Ballew & Todorov, 2007; Spezio et al., 2008). The most notable example are smiles, which are known to increase judgments regarding attractiveness (Golle et al., 2013; Mueser et al., 1984; O'Doherty et al., 2003), trustworthiness (Centorrino et al., 2015; Krumhuber et al., 2007; Schmidt et al., 2012), and competence (Min & Hu, 2022; Wang et al., 2017). Smiles and attractiveness have been observed to translate into potential benefits in laboratory settings (Andreoni & Petrie, 2008; Póvoa et al., 2020), in online peer-to-peer markets (Jaeger et al., 2019), as well as in remote, indigenous societies (Godoy et al., 2005).

Within the face, specific attention is given to the eyes. The large proportion of white sclera to dark central iris allows humans to learn about the environment and others through gaze attention, joint and shared attention (Emery, 2000). A well-known measure regarding the ability to infer the psychological states of others (i.e., the "Reading the Mind in the Eyes" test; see (Baron-Cohen et al., 2001), uses photographs of the eye region and asks participants to choose from a set of words what the person in the photo is feeling. Face masks thus leave intact one of the main channels, but not the only one, for detecting emotional expressions in others. For instance, in many Western cultures, the mouth is a much more expressive part of the face than the eyes, and attention is directed at it (Yuki et al., 2007). Furthermore, recent studies argue that when others wear masks, the intensity of perceived emotions is diminished (Marini et al., 2021; Tsantani et al., 2022), as are ratings regarding happiness (Biermann et al., 2021) and the recognition of positive and negative emotions (Pavlova & Sokolov, 2022).

Given this impairment concerning the detection of subtle facial signals on a masked face, we might expect that individuals are less influenced by unjustified biases with respect to personality judgments and halo effects. However, if face masks center our attention on the eyes and make individuals more aware of them, individuals may still use these informational cues to judge others. The aim of this paper is thus to investigate to which degree, given reduced abilities to convey non-verbal signals with the face, such signals will influence the ratings and judgments of others. We focused on the most prominent non-verbal facial signal, smiles, and investigated ratings regarding the face (i.e., attractiveness) but also regarding the individuals' personality (i.e., trustworthiness and competence). Imposed—or self-imposed—mask-wearing mandates are especially likely to influence how we interact with strangers. That is, individuals of whom we have never seen their faces. Therefore, we focused on studying reactions to different, unknown faces covered by face masks.

We investigated the impact of face masks through three complementary studies. Study 1 established that smiling faces are easily distinguished from neutral faces even when covered by a face mask. In Study 2, we presented individuals with smiling and neutral faces to investigate whether ratings regarding attractiveness, trustworthiness, and competence are influenced by smiles, even if the face is covered with a face mask. Finally, in Study 3, we studied individuals' awareness of influencing perceptions by others when wearing a face mask. This is important because face masks are likely to become less of an obligation and more of a recommendation over the near future. Thus, individuals' perceptions of how face masks influence their interactions with others are likely to influence the use and acceptance of masks. Study 3 thus presented participants with masked smiling or neutral faces and asked individuals to predict ratings of these faces regarding attractiveness, trustworthiness, and competence. In addition, individuals were asked to imagine how changing a smiling expression to a neutral expression, or changing a neutral expression to a smiling expression, would influence these ratings. Since all faces were previously rated in Study 2, these predictions can be compared to actual ratings.

Our results showed that smiling faces are easily distinguished from neutral expressions even when covered by face masks. Smiling behind face masks also significantly and positively affects ratings regarding attractiveness, trustworthiness, and competence. Finally, the effect regarding attractiveness and trustworthiness seems to be anticipated by our participants, that correctly estimated that these traits are rated higher for a smiling person even if a face mask covers the smile. By contrast, participants did not seem to anticipate how smiling behind a mask affects competence. Face masks thus do not seem to change how we evaluate others when it comes to basic signals such as intentional smiles. It is important to note that smiles were not necessarily genuine and reflect the situation individuals experience when deciding whether to smile for a photograph or stranger. Our results thus provide a lower bound of the effects since more genuine expressions, reflecting the experience of spontaneous happiness, might lead to even stronger effects.

2. Study 1: Distinguishing smiling from neutral masked faces

When investigating the effect of smiles under face masks, a first step is to verify that smiles can be detected even when covered by a mask. Previous studies investigated the accuracy when rating different emotions when covered by face masks and showed that noise in ratings is increasing (Pavlova & Sokolov, 2022). This result is not surprising, given that the mask covers a large part of the face. However, our question was much simpler. We did not want to compare accuracy to ratings of unmasked faces but observe the ease with which smiles can be distinguished from neutral expressions.

2.1. Methods

Thirty-five photo volunteers (17 female; 18 male) were contacted to provide two photographs of themselves on which they were asked to wear a mask covering their mouth and nose. Most of the photo volunteers were former students (14), colleagues (8), close relatives from the research team (4), and other volunteers recruited by a research assistant (9, primarily undergraduate students). Most

of the photo volunteers were somehow related to higher education. On one of the otherwise identical photographs, they were instructed to smile. In contrast, they were asked to look neutral on the other. Volunteers also provided smiling and neutral photographs without a mask for reference. However, in all three studies, we only employed the masked photographs. Photo volunteers provide consent to use the photographs in the study.

In the study, we used the photographs from thirty photo volunteers, with a 50–50 sex ratio. Two of the other five were discarded because the volunteers were under-aged, and another did not sign the informed consent. We further used the photographs of two participants as illustration material in the study instructions because the smiles were very evident.

To understand the differentiation of smiling from neutral masked faces, we presented these photographs to participants in an online study consisting of two parts. In Part 1, participants were presented with ten pairs of photographs. In each pair, the individual in one photograph smiled while the other individual showed a neutral face. Both were of the same sex. Participants were asked to indicate the smiling individual for each pair. Fig. 1 offers an idea of how photographs were displayed in the task. Responses were collected in two treatment conditions regarding exposure time. In the ‘long exposure’ treatment, participants had 30 s to look at the photographs and submit their choices. In the ‘short exposure’ treatment, participants had only 5 s.

We programmed and conducted the study in LimeSurvey. In Part 1, participants saw two photographs and were asked first to select (i.e., click) the photograph they considered as smiling and then press a confirmation button. Participants then saw a blank page with a “Next” button. They were informed that the next stimulus would appear once they pressed this button, and the time counter, displayed on top of the page, was reset. If time ran out, both photographs disappeared from the page. If participants had already selected a photograph by that time, their response was recorded, even if they had not yet activated the confirmation button. If no photograph was chosen when time ran out, participants received a null score for this question (i.e., it was counted as an incorrect answer), and the participant proceeded to the next blank page mentioned above. In the second part, participants were presented with an additional set of ten photographs of individuals that they had not previously seen. They were asked to rate the person in the photograph regarding attractiveness and trustworthiness (see Online Appendix F). All the photographs in the second part displayed an individual with a neutral expression.

The experiment ended with a short sociodemographic survey regarding gender, age, country of residence, and everyday exposure to face masks. We also asked participants if any of the thirty faces they had seen looked familiar. Specifically, we asked: “Do you know, or think you know, any of the people you just saw during the test?”.

We advertised this study on social media and through the institutional network of the researchers. Participants entered through a link to LimeSurvey, where the study was conducted. No financial incentives were involved. At the end of the experiment, we informed participants about their score from the first part, next to an invitation for their friends to take the test.

We created three different exposure sets, A, B, and C, to ensure that participants never saw the same volunteer twice and that all photos received the same number of ratings (see Online Appendix A for details). Participants were randomly assigned to one of these exposure sets. For example, volunteers whose smiling picture was added to set A had their neutral picture displayed in set B and were added to the personality assessment in set C. All the participants assigned to a specific exposure set received the same questions. This matching procedure guaranteed that we have three different photograph matchings to be assessed in Part 1, and that all the photographs were assessed in Part 2, though by three different groups of respondents, depending on the assigned exposure set.

2.2. Transparency and openness

We described our sampling plan, all data exclusions (if any), all manipulations, and all measures in the study, and we adhered to the *Journal of Applied Psychology* methodological checklist. Data were analyzed using Stata 17 (StataCorp, 2021). This study’s design and its analysis were not preregistered.

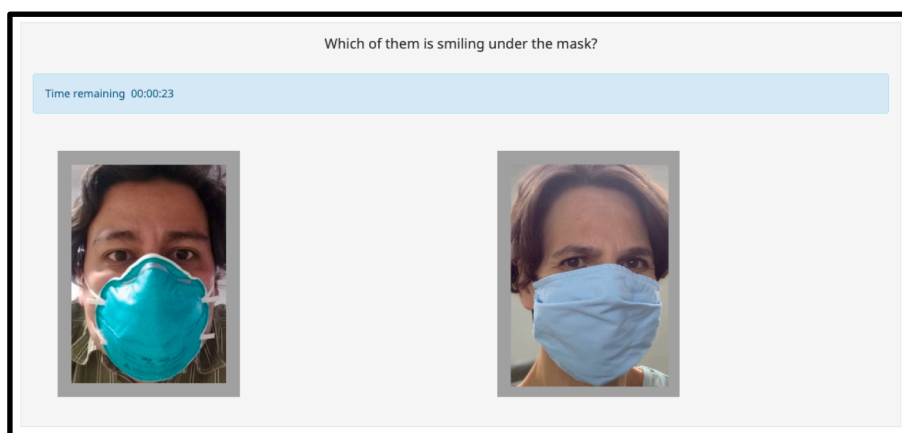


Fig. 1. Illustration of the task in Study 1. The original photographs were replaced by the authors’ photographs for data protection reasons. Note that in the study both individuals displayed were of the same sex.

2.3. Data

A total of 1,814 individuals completed the questionnaire between October 2020 and February 2021. We excluded 52 observations from underaged respondents and 38 observations from participants that took the test more than once, keeping only their first test. Approximately 99% of the observations were collected in the first two months. Participants took an average of 9.3 min (median 6.4) to complete the survey. The completion time by time exposure treatment is not significantly different (two-tailed *t*-test, $p = 0.916$). Participants could take the questionnaire from mobile devices or computers. The majority of responses came from smartphones (70.5%), followed by computers (28.5%) and tablets (1%). Roughly two-thirds of the respondents self-identified as female (64.8%), whereas 34% self-identified as male, and 1.2% as non-binary. Participants were, on average, 28.8 years old (std. dev. 10.3).

The questionnaire was available in Spanish, English, and French. It was disproportionately taken in Spanish (95%), followed by English (4.5%). This can be explained by the use of the authors' institutional networks to disseminate the questionnaire. The country with the most extensive participation in our test was Colombia (80.5%), followed by France (3.2%), the United States (2.9%), Germany (2.1%), Mexico (1.8%), and Denmark (1.3%). Thirty-two percent of participants reported familiarity with at least one of the shown faces. Among those replying positively, the median number of familiar faces was 3 (average 5.3).

Participants reported that, on average, they wore masks 95% of the time in public spaces (with 75% of them reporting that they wore them one hundred percent of the time). The main reason to wear a mask was to protect themselves (36.8%), followed by the desire to protect their beloved ones (32.7%) and to protect others (19.7%). Only 10.8% of the respondents listed that it was mandatory, as their main reason to wear it. When asked to estimate the percentage of other people wearing a mask in public spaces, participants reported an average of 81%.

2.4. Results

2.4.1. Part 1: Discrimination between smiling and neutral faces

Participants were exposed to a set of ten pairs of pictures. Regarding their ability to discriminate smiling from neutral masked faces, they could therefore obtain a score ranging from 0 to 10. The overall average score of participants was 8.7 (std. dev. 1.26), with 28% of participants obtaining a score of 10 and another 38% obtaining a score of 9.

A two-tailed *t*-test showed that the score was significantly higher in the *long exposure* treatment compared to the *short exposure* treatment (9.00 versus 8.33; $p < 0.001$). In Online Appendix B, we discuss the balance between exposure sets. We show in Table C.1

Table 1
Distinguishing smiling from neutral faces (0 to 10) [Study 1].

Dependent variable: Score (out of 10) in the detection of smiling faces	(1)	(2)	(3)
Long exposure	0.479*** (0.105)	0.474*** (0.103)	0.416*** (0.102)
Device: Mobile	-0.220* (0.090)	-0.233* (0.091)	-0.242** (0.091)
Long exposure × Mobile	0.283* (0.124)	0.268* (0.122)	0.317** (0.121)
Total number of familiar faces (self-identified as) female		0.027*** (0.008)	0.028*** (0.008)
Age		0.035 (0.059)	0.017 (0.059)
Language: Spanish		-0.015*** (0.003)	-0.015*** (0.003)
Country: Colombia		-0.190 (0.127)	-0.053 (0.145)
Frequency for own use of mask			-0.107 (0.084)
Frequency for others' use of mask			-0.003 (0.003)
Reason for wearing mask: Mandatory			0.003 (0.002)
Reason for wearing mask: Protect others			0.033 (0.099)
Reason for wearing mask: Protect beloved ones			0.076 (0.079)
Constant	8.588*** (0.086)	9.172*** (0.168)	9.236*** (0.259)
Observations	1,814	1,739	1,644
R-squared	0.122	0.147	0.154

Notes: Ordinary least squares regression. Other covariates included in the regression: categorical variables for exposure sets B and C. Standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

(see Online Appendix C) that, under long exposure, the distance to a perfect score is mostly explained (86%) by mistakes. By contrast, under short exposure, it is almost equally explained by mistakes (48%) and unanswered stimuli (52%).

We complemented this finding with a regression analysis reported in Table 1. The ordinary least squares – OLS regression results confirmed that, controlling for other covariates, a long exposure (i.e., 30 s) to the stimuli increased the score by approximately 0.48 units when using a computer and by 0.7 units for mobile devices (i.e., smartphones or tablets). The reason for this difference is that pictures were displayed side-by-side on large computer screens but required some scrolling on mobile devices. Thus, it is unsurprising that the participants who benefited the most from relaxing the time constraint were those responding to the questionnaire from a mobile device.

We further found that reporting to be familiar with more faces increased the predicted score. Nonetheless, the effect is not very large: being familiar with half of the faces would have a slightly smaller effect (0.41 units) than the long exposure of each pair of faces. Notethat, since the coefficients of language and country are statistically non-significant, we believe that the reported familiarity is not necessarily associated with the cultural similarity of facial traits. We also observed a small but significant coefficient for the age of participants. We observed no significant differences regarding the participant's gender, self-reported frequency of mask use by themselves or others, or the most important reasons for mask use.

Higher familiarity with the observed faces may be due to some unobservable commonality of these participants. Thus, the existence of some self-selection in our study. We verified the robustness of our results by running additional regressions that excluded participants reporting at least one face that looked familiar (559 respondents). The results, reported in Online Appendix C, Table C.2, are qualitatively identical to those reported in Table 1.

A potential concern regarding the detection of smiling faces is that, depending on the type and shape of the face mask, different facial areas are visible and thus provide more or less information about the facial expression. We report in Online Appendix C, Table C.3, that neither the mask type (i.e., cloth or surgical), the amount of unexposed area, nor their color, were correlated with detection rates.

2.4.2. Part 2: Attractiveness and trustworthiness ratings of neutral faces

Ratings regarding attractiveness and trustworthiness for men and women are displayed in Figure D.1. Average ratings regarding attractiveness were 5.1, with average ratings per photograph ranging from 3.7 to 7.2. Photographs of women were rated on average as more attractive than men (t -test, $p = 0.003$ with $N = 30$), with mean ratings for male photographs ranging from 3.8 to 5.4 and for female photographs from 3.9 to 7.2.

Average ratings regarding trustworthiness were 5.9, with average ratings per photograph ranging from 4.2 to 7.2. Again, photographs of women were rated on average as more trustworthy than men (t -test, $p < 0.001$ with $N = 30$), with mean ratings for male photographs ranging from 3.8 to 5.4 and for female photographs from 3.9 to 7.2. See Figure E.1, in the Online Appendix, regarding means and standard deviation of ratings for all photographs.

Finally, ratings regarding trustworthiness and attractiveness were also not related to the masks' type, shape, color, or the amount of unexposed area in the face (see Table C.3).

2.5. Discussion

Our results clearly indicated that smiles are easily distinguished from neutral expressions, even if a face mask covers the face. Time exposure, as measured by our study (5 versus 30 s), did not influence the outcome to a large extent. Given this apparent ease with which smiles were detected even from people wearing face masks, we thus wondered to which degree the detection of smiles influences ratings regarding traits of others. More specifically, we wanted to understand to which degree ratings regarding attractiveness and trustworthiness would differ for faces that were either smiling or neutral. Moreover, knowing that face familiarity increases the ability to detect smiles behind masks, the effect of smiles on the ratings of attractiveness and trustworthiness is much more informative if the rated photographs belonged to strangers. We thus focused in Study 2 on participants who were unlikely to know any of the photo volunteers. This way, we can be confident that ratings are not affected by other uncontrolled sources of information related to the photo volunteers. In addition, participants in Study 1 were allowed to respond from different types of devices, which enabled us to reach a large audience (more than 1,800 participants). However, we found a slightly lower performance among smartphone users that were time constrained. To avoid any confound, in the next study we thus decided to ask participants to use computers.

In Study 2, we compared ratings made in a more controlled environment (i.e., recommending the use of computers and providing incentives) with respect to smiling and neutral faces wearing a mask, from participants for whom it is extremely unlikely that they knew any of the photo volunteers.

3. Study 2: The impact of smiles behind face masks on ratings regarding attractiveness, trustworthiness, and competence

Study 1 provided clear evidence that smiles can be easily and quickly detected, even for individuals wearing a face mask. For unmasked faces, smiling is known to influence personality judgments in an important way: notably, smiling faces are considered to be more attractive (Golle et al., 2013; Mueser et al., 1984; O'Doherty et al., 2003) and are rated as more trustworthy (Centorrino et al., 2015; Krumhuber et al., 2007; Schmidt et al., 2012). Thus, we wondered whether this bias also carries over to smiles covered by a face mask. In the previous study, we collected ratings of attractiveness and trustworthiness of faces displaying a neutral expression. In Study 2, we investigated whether these ratings are affected by smiling. We also expanded the studied judgments to include competence, another trait judgment influenced by smiles (Min & Hu, 2022).

While ratings regarding attractiveness represent an individual's personal taste and evaluations of a face, ratings regarding trustworthiness attempt to predict "something" about the observed person (e.g., infer her behavior). Trustworthiness judgments are very much influenced by how familiar we perceive others (Lavan et al., 2021) and how much we like someone (Giang et al., 2012). They are also highly correlated with attractiveness (Todorov, 2008).

An even further step is taken when rating competence because it does not only attempt to judge personality but also knowledge and intelligence (Fiske et al., 2007). We thus expected that the halo effect of smiling would be strongest for ratings regarding attractiveness and trustworthiness and weakest for ratings regarding competence.

3.1. Methods

To simplify the random exposure to photographs, we selected a subset of twenty-four photo volunteers from the thirty employed in Study 1. For this, we excluded the three male and three female photo volunteers for which the correct identification of the smiling faces was too evident.

The resulting smiling and neutral photographs were presented to a new set of participants recruited through the online platform Prolific. To control for the ability to discriminate between smiling and neutral expressions, we replicated Part 1 from the previous study and asked participants to indicate the smiling photograph for six pairs of pictures of different individuals. Each participant faced three pairs corresponding to female and three to male photographs. This replication was helpful for three reasons. First, the English language Prolific sample eliminated any issue related to familiarity. Second, each smiling face was randomly matched to one of three different neutral faces. In the first study, a given smiling face was always matched with the same neutral face from another participant. The random configuration of stimuli in Study 2 reduced concerns that smile detection was explained by joint properties of the matched faces and not by individual properties of the smiling photograph. Third, we introduced monetary incentives to participants in the study by paying for accuracy in Part 1. In Part 2, participants were presented with the remaining set of twelve photographs (6 male and 6 female) and asked to rate the person on the photograph regarding attractiveness, trustworthiness, and competence (see Online Appendix G for the full instructions of Study 2). Given our interest in the difference in judgments from masked faces when individuals are smiling or not, each participant was presented with half of these photographs from volunteers that were smiling, and the other half from volunteers that were neutral. The smiling or neutral version was randomized for each participant. This feature ensured that nobody saw more than one image of the same person.

Photographs were divided into two groups. Participants were exposed to one of these groups in Part 1 and the other in Part 2. This grouping of stimuli yielded two exposure sets, D and E, different from the sets A, B, and C employed in Study 1. In each exposure set, participants responded in Part 1 with respect to one group of photographs and were asked to rate the other twelve photographs in Part 2. Further details regarding the construction of stimuli can be found in Online Appendix A. We also included a short sociodemographic survey at the end of the experiment, including gender, age, country of residence, and everyday exposure to face masks. Table B.2 in the Online Appendix reports the balance check for participants' characteristics across the two sets of stimuli.

As mentioned above, the study was conducted on the online platform Prolific. Participants received a participation fee of £1.00 for completing Parts 1 and 2, plus £0.10 for each photograph pair correctly identified in Part 1 (i.e., which one was smiling). Given that the score differences between the 'long exposure' and 'short exposure' treatments in Study 1 were partially explained by the device from which the participant took the test, we fixed the available time per photograph pair at 30 s and strongly recommended the use of computers.

We conducted an OLS regression with the assessed trait as the dependent variable, with the main explanatory variable being the photograph's smiling (or neutral) look. This multivariate model allowed us to include two types of covariates aimed at reducing potential confounding effects. First, we included one categorical variable for each photograph volunteer (i.e., fixed effects coefficients per photograph volunteer, in the multilevel regression terminology) in the regressions. They captured any specific volunteers' characteristics not directly included in our multivariate model (e.g., expressiveness of a face or the type of face mask employed).

Second, we also included one categorical variable for each respondent (i.e., fixed effects coefficients per respondent) to capture unobserved differences among the participants in our experiment (e.g., attention to the task, or extremeness of the respondents' overall assessment of traits). These categorical variables are equivalent to having one intercept per respondent, capturing the *average* rating in her judgment of faces. Hence, our coefficient of interest, telling us whether smiling faces are judged more positively, is robust to alternative explanations where higher ratings for smiling faces are due to unobservable differences of raters.

3.2. Transparency and openness

We described our sampling plan, all data exclusions (if any), all manipulations, and all measures in the study, and we adhered to the *Journal of Applied Psychology* methodological checklist. Data were analyzed using Stata 17. This study's design and its analysis were not preregistered.

3.3. Data

Two hundred and fifty individuals participated in this study between June and July 2021. Participants took, on average, 6.8 min (median 6.1 min) to complete the survey. The majority of participants responded from computers (79.2%), with the remainder using smartphones (18%) and tablets (2.8%). Participants were, on average, 27.0 years old (std. dev. 8.6) and 46.4% self-identified as female (53.2% as male). The questionnaire was available only in English. Most of the respondents reside in the United Kingdom (23.5%),

followed by Portugal (16.2%), South Africa (11.3%), Poland (10.5%), Greece (6.9%), and Italy (6.1%).

Participants reported wearing a mask in public spaces on average 80.6% of the time (with 44% reporting wearing a mask all the time). The main reason to wear a mask was its mandatory character (32.5%), followed by the desire to protect themselves (29.2%) and their beloved ones (20.2%). Participants perceive that, on average, other people wear masks in public spaces about 67.5% of the time.

3.4. Results

3.4.1. Part 1

The average score regarding smiling face detection was 4.7 (out of six, std. dev. 0.89, median score: 5). Eighteen percent of participants detected the smiling face in all six cases. These results confirm that smiles behind a face mask are easily detectable. The normalized score ($4.7/6 = 0.78$) is slightly lower than in Study 1 (0.87). However, this can be explained by our discarding of the six photo volunteers whose smiling photograph was more easily detectable in Study 1.

3.4.2. Part 2

Faces were rated with respect to attractiveness, trustworthiness, and competence (on a scale from 1 to 10). While attractiveness is a subjective rating based on facial perception, trustworthiness and competence ratings aim to predict a *de facto* unobservable characteristic and thus are potentially influenced by halo effects (Forgas & Laham, 2016).

The average attractiveness rating was 5.12 (std. dev. 2.27). Smiling faces were rated as significantly more attractive (5.32) than neutral ones (4.91) (*t*-test $p < 0.001$). This result also holds on the individual's face stimuli level: 18 of 24 faces were rated as more attractive when the person smiled (and 10/18 of the *t*-tests are statistically significant). Thus, even though the exposed facial area is much reduced when wearing a face mask, signals of smiles detected in these faces lead to higher attractiveness ratings.

We further observed that ratings with respect to trustworthiness and competence were also significantly higher for smiling faces than for neutral faces. Average ratings regarding trustworthiness were 5.91 (std. dev. 2.03), with smiling faces receiving an average rating of 6.17 and neutral photographs a rating of 5.65 (two-tailed *t*-test, $p < 0.001$). Also, most of the volunteers' faces were rated higher on trustworthiness when the face was smiling (20 out of 24 faces; and 13/20 of the *t*-tests were statistically significant).

Average ratings for competence were 6.14 (std. dev. 1.95), with a significant difference between smiling (6.23) and neutral faces (6.03; *t*-test with $p = 0.004$). Again, the result also holds on the volunteer's face level, with 16 out of 24 faces rated as more competent when smiling (and 9/16 of the *t*-tests were statistically significant).

The overall increase in ratings regarding all three dimensions for smiling photographs can also be seen in Fig. 2. We complemented these results with regressions controlling for participant and photograph fixed effects coefficients (see Table 2). We observed that attractiveness ratings increased by about 0.4 points when comparing smiling to neutral faces (i.e., an increase of around 8%). Ratings concerning trustworthiness increased even more: smiling faces are rated 0.6 points (approximately 9%) higher. Ratings concerning competence increased by an average of 0.2 points (approximately 3%). These results also hold for roughly two-thirds of the respondents when computing separate regression analyses for each of them (see Figure E.2 in the Online Appendix).

3.5. Discussion

Study 2 confirmed that smiles are easily distinguished from neutral faces when wearing a face mask. Even though mostly covered by a mask, smiling has a strong halo effect that carries over to ratings regarding attractiveness, trustworthiness, and competence. The

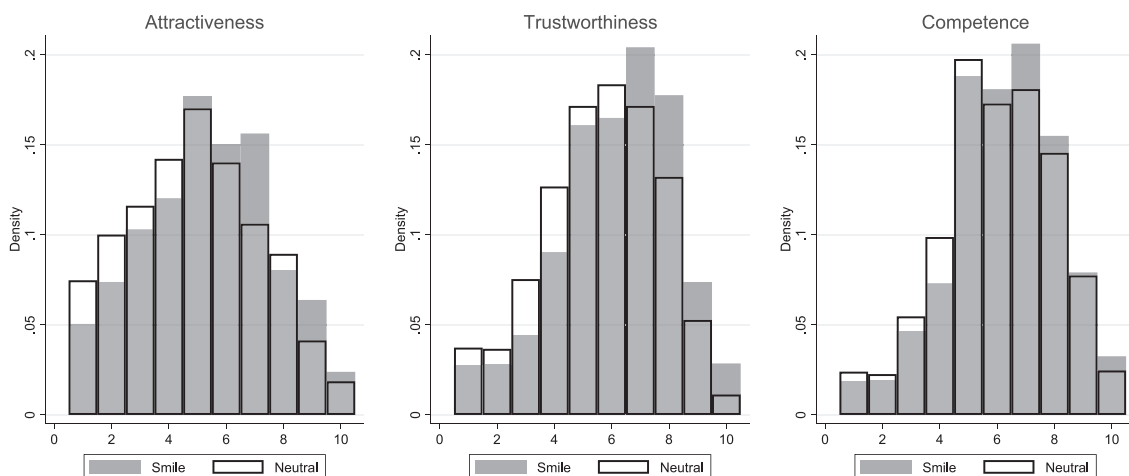


Fig. 2. Ratings of smiling versus neutral masked faces [Study 2].

Table 2
Ratings regarding attractiveness, trustworthiness and competence dependent on smile [Study 2].

Dependent variable: Rating [1–10 scale]	Attractiveness			Trustworthiness			Competence		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Smile	0.422*** (0.0909)	0.398*** (0.0828)	0.397*** (0.0689)	0.517*** (0.0814)	0.533*** (0.0790)	0.620*** (0.0746)	0.205** (0.0742)	0.224** (0.0718)	0.260*** (0.0640)
Constant	4.415*** (0.628)			6.004*** (0.551)			5.753*** (0.570)		
Photograph FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Respondent FE	No	No	Yes	No	No	Yes	No	No	Yes
N	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000	3,000
R-squared	0.012	0.262	0.602	0.018	0.129	0.465	0.005	0.113	0.479

Notes: Other covariates included in the regression: categorical variables for exposure sets B and C, and score in Part 1 (recognition of smiles). We omit the intercept in models including fixed effects coefficients for clarity. Standard errors clustered at the respondent level are shown in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

effect is strongest regarding ratings of trustworthiness. This result aligns with (Ert & Fleischer, 2020) exploration of various photo characteristics in an applied setting and the stronger correlation between smiling and trustworthiness than between smiling and attractiveness.

Given this observed effect of smiles, we might wonder to which degree individuals are aware of the impact that smiling has on ratings regarding attractiveness, trustworthiness, and competence. In Study 3, we thus asked individuals to make predictions regarding judgments by others.

4. Study 3: Assessing beliefs regarding the impact of smiles behind masks on perceptions by others

The first two studies showed that smiles could be easily detected in individuals wearing face masks. Moreover, like in the case of uncovered faces, smiles strongly influence perceptions regarding attractiveness, trustworthiness, and competence. Though the widespread use of face masks is still somewhat recent, we thus wondered to which degree individuals are aware of these effects. We ran a third study in the fall of 2021—about 18 months after the start of the pandemic, to investigate beliefs about how smiling under a face mask influences judgments by others.

4.1. Methods

For Study 3, we designed an experiment in which participants were asked to predict the average ratings of attractiveness, trustworthiness, and competence collected in Study 2. We kept a subset of eighteen photographs, excluding photographs whose ratings were close to the average on the three dimensions. In addition to submitting predictions regarding ratings by others of the photographs they were seeing (either smiling or neutral), participants were further asked to make predictions about ratings for the counterfactual state. That is, when seeing a smiling photograph, they were asked to predict how the same person had been rated when looking neutral, and the other way around. This design feature allows us to measure beliefs regarding the impact that smiling has on perceptions by others (in addition to the general accuracy of predictions).

Participants were presented with eighteen photographs and asked to predict ratings regarding attractiveness, trustworthiness, or competence. We presented six photographs for each trait, three from female and three from male participants (for details on the construction of stimuli, see Online Appendix A). We explained to respondents that, in a previous study, participants rated the smiling and neutral photographs of the people shown to them. Respondents in Study 3 were shown only one of the two photographs, and they knew whether they were seeing the smiling or neutral version. Then, they were asked to submit the prediction regarding the average rating of the picture they were seeing, and regarding the average rating of the picture they were *not* seeing. If they were informed that a given picture corresponded to the smiling version, they were asked to predict the average rating of the corresponding trait (attractiveness, trustworthiness, or competence) for this smiling photograph, and also for the neutral and unseen photograph of the same person.

In Study 2, participants rated photographs with any integer between 1 and 10. However, the dependent variable in Study 3, predictions of the mean ratings, had to be continuous. To avoid bunching of responses at integer values we asked participants in Study 3 to make predictions on a scale from 10 to 100. Predictions were rescaled and compared to average ratings from Study 2 (up to one decimal behind the comma), which we refer to as the “true values.” Participants were recruited and paid through Prolific and received a baseline payment of £1.00 for completing the study, plus bonuses based on the accuracy of predictions. Each prediction within a range of 0.5 points from the true value gave £0.10 extra, and each prediction within a range of up to 1 point from the true value gave a £0.05 bonus. Participants were free to take as much time as they wanted, though the task was announced to take about 10 min. Participants also filled out a short sociodemographic questionnaire at the end of the survey (see Online Appendix H for the full instructions of Study 3).

We first investigated the accuracy of predictions for smiling and neutral and for observed and counterfactual faces. To study the expected effects of a smile, we exploit the fact that participants predicted each face twice (smiling and neutral), regardless of which of the two photographs was displayed. Our dependent variable is the difference between the predicted ratings regarding attractiveness,

trustworthiness, and competence for the smiling versus the neutral photograph. The intercept in an OLS regression can be used to understand whether the impact of smiles under face masks is over or underestimated. Accuracy in predicting the impact of smiles might further be influenced by the information available to the participant. That is, it might be harder to imagine the effect of a smile (when not seeing one) than the effect of a neutral expression (when seeing a smile). We thus introduced as an explanatory variable whether the observed photograph was the smiling photograph or the neutral one. The regressions further include face and respondent fixed effects coefficients.

We present in Online Appendix D additional regression exercises for each of the rated judgments that are directly explained in the Online Appendix. First, we explored the effect that smiles and the type of observed photograph, have on the predicted ratings. Second, we explored if the same explanatory variables that we identified in Study 2 affect the accuracy of predictions in Study 3.

4.2. Transparency and openness

We described our sampling plan, all data exclusions (if any), all manipulations, and all measures in the study, and we adhered to the *Journal of Applied Psychology* methodological checklist. Data were analyzed using Stata 17. This study's design and its analysis were not preregistered.

4.3. Data

Ninety-four individuals were recruited through Prolific to participate in this study in November 2021. The survey was completed in 11.4 min on average (median 10.2). Most participants responded from computers (71.3%), with the remainder using smartphones (25.5%) and tablets (3.2%). Participants were, on average, 25.8 years old (std. dev. 7.3), and 53% self-identified as female (45% as male). The survey was conducted in English.

Participants reported a mean frequency of mask use in public spaces of 85% (with 52% reporting wearing it at least 99% of the time). The main reason to wear a mask was to protect themselves (38.2%), followed by its mandatory character (28.1%) and the desire to protect their beloved ones (19.1%) and the rest of the population (14.6%). Participants perceived that others wear masks in public spaces, on average, 74.2% of the time. These values are similar to those observed in Study 2.

4.4. Results

To compare predictions (scale of 10 to 100) to ratings from Study 2 (scale of 1 to 10), we report adjusted predictions (i.e., predictions divided by 10). The average adjusted prediction for attractiveness rating was 5.36 (std. dev. 2.51), a higher value than the average attractiveness ratings in Study 2, of 5.12 (a one-sample two-tailed t -test yields a $p < 0.001$). Predictions of attractiveness ratings were higher for smiling (5.75) compared to neutral faces (4.98; two-tailed t -test, $p < 0.001$). The difference in the attractiveness rating between observed (5.23) and unobserved (5.50) faces was not statistically significant (two-tailed t -test $p = 0.075$).

The average adjusted prediction regarding trustworthiness ratings was 5.83 (std. dev. 2.38), not different from the average of 5.91 in Study 2 (one-sample two-tailed t -test yields a $p = 0.265$). Predictions of trustworthiness ratings were higher for smiling (6.15) compared to neutral (5.51) faces (two-tailed t -test, $p < 0.001$). The difference between predictions for observed (5.74) and unobserved (5.91) photographs was not statistically significant (two-tailed t -test, $p = 0.244$).

The average adjusted prediction for the competence rating was 5.97 (std. dev. 2.30), a lower value than the average competence rating in Study 2, 6.14 (one-sample two-tailed t -test yields a $p = 0.012$). The difference in predictions between smiling (6.03) and neutral (5.90) faces was not statistically significant ($p = 0.363$). The differences between observed (5.91) and unobserved (6.02) faces were not statistically significant either.

($p = 0.425$).

Table 3 reports results from OLS regressions explaining the difference in predicted ratings regarding attractiveness (columns 1, 2),

Table 3

Predictions regarding ratings of attractiveness, trustworthiness and competence dependent on available information [Study 3].

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Attractiveness		Trustworthiness		Competence	
Observed smile	-0.533** (0.169)	-0.517** (0.184)	-0.331 (0.242)	-0.279 (0.279)	-0.219 (0.219)	-0.180 (0.227)
Constant	1.042*** (0.146)		0.816*** (0.215)		0.234 (0.176)	
Linear test:						
Constant + Obs. Smile	0.509*** (0.108)		0.485*** (0.117)		0.015 (0.142)	
Face fixed effects	No	Yes	No	Yes	No	Yes
Respondent fixed effects	No	Yes	No	Yes	No	Yes
Observations	564	564	564	564	564	564
R-squared	0.022	0.314	0.005	0.306	0.002	0.298

Notes: We omit the intercept in models including fixed effects coefficients for clarity. Standard errors clustered at the respondent level are shown in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

trustworthiness (columns 3, 4), and competence (columns 5, 6) between smiling and neutral faces. We first focus on the accuracy of predictions, independent of which version of the face was shown. We turn our attention to the regressions' intercept in the odd columns, where we found a positive constant in the baseline model for the three predicted ratings (1, 3 and 5).

The coefficient is strongly significant in the case of attractiveness ratings. This finding implies that participants correctly understood that smiling faces were rated more attractive than neutral faces. The coefficient size can be compared to actual ratings to see if predictions were over- or underestimating the effect of a smile. Regarding attractiveness, recall that average ratings of smiling faces in Study 2 were approximately 0.4 points higher than for neutral faces. The magnitude of the constant reported in Table 3 suggests an overestimation of the effect of smiles.

The effect for trustworthiness ratings is similar. Participants expected that this rating increased by about 0.8 points for smiling faces, which is overestimated by about 0.5 points with respect to the effect reported in Table 2. Finally, the coefficient regarding competence is not significant though being positive (0.23, with a standard error of 0.18). This result implies that participants did not seem to anticipate that smiling photographs induce higher ratings in terms of competence.

To see whether the predicted effect of smiling is influenced by the visual information available to the participant, we look at the coefficients of the variable indicating if participants saw the smiling version of the photograph. Recall that the effect of smiling was overestimated for the attractiveness rating. This overestimation was almost fully corrected when participants saw the smiling photograph (cf. column 1, net effect $1.042 - 0.533 = 0.509$). The estimated coefficients are robust to the inclusion of fixed effects. Thus, participants seemed to overestimate the effect of smiling on attractiveness ratings when observing the neutral photograph but were rather precise in estimating the increase in attractiveness when observing the smiling version.

A similar trend can be observed regarding predictions of trustworthiness ratings, although the effect of seeing the smiling photograph is not statistically significant. The constant term suggests that participants slightly overestimated the effect of the smile. However, when participants observed the smiling photograph, the predicted effect of a smile fell to $(0.816 - 0.331 =) 0.485$, a value close to the effect found in Study 2. For competence, there seemed to be a much smaller effect when observing the neutral photograph, which was almost zero ($0.234 - 0.219 = 0.015$) when observing the smiling photograph. In other words, participants from Study 3 were aware that smiles improve trustworthiness ratings, but they seemed to neglect, or at least underestimate, the effect of a smile on competence ratings. Nevertheless, their predictions were more precise when seeing the actual smile.

4.5. Discussion

The results from Study 3 showed that individuals are clearly aware of the positive effect of smiling on attractiveness ratings, even for people wearing face masks. They tended to overestimate the impact when not actually seeing the smiling face.

On average, individuals also seemed to be aware of the effect of smiling on trustworthiness ratings. However, the observed photograph (i.e., smiling or neutral) was less important in predicting the effect of smiles. Finally, our results showed that while smiling under masks also influenced perceptions regarding competence, this effect was not anticipated by the participants in Study 3.

5. Conclusions

In many countries around the globe, mask mandates are currently relaxing and giving more freedom to individuals to choose whether to wear masks. Hence, studying the perceived impact of wearing face masks on perceptions and interactions with others becomes ever more important. In this paper, we studied the capacity to distinguish smiling from neutral faces and how smiling under a mask influences perceptions regarding the wearers' attractiveness, trustworthiness, and competence. While these attributes are known to be influenced by smiling when individuals are unmasked, we have yet to study whether mask-wearing would alleviate these biases.

Our results show that, even for masked faces, people have minor problems distinguishing smiling from neutral expressions. Even though face masks cover up a large proportion of the face, smiles lead to higher ratings regarding attractiveness, trustworthiness, and competence. Moreover, individuals seem to be aware of how smiling affects others' judgments regarding attractiveness and trustworthiness. Our results thus provide evidence that face masks do not seem to impact the already-known fact that smiles influence how others perceive us. About 18 months of experience with face masks seem to have been enough for individuals to be aware of this and to appreciate that, even under masks, smiles are detected in others.

Theoretically, our results show that evaluations regarding attractiveness, trustworthiness, and competence are biased by a smile, even if given very little information. While an extensive literature has previously investigated that such impressions are formed very rapidly, our results show that very little information available from the face (i.e., only the eye region) is sufficient to form beliefs. The similar effect on attractiveness and trustworthiness ratings also points to the importance of linking the literature regarding the beauty premium (e.g., (Andreoni & Petrie, 2008)) to determinants of trust (e.g., (Bejarano et al., 2021; Bonowski & Minnameier, 2022; Sofianos, 2022)).

Initial concerns among the general public conveyed the idea that face masks might inhibit everyday non-verbal communication. Our results imply that these hesitations should be largely eradicated by now. Our study intentionally focused on typical participants smiling or not for a photograph and is thus representative of the situation individuals face in everyday life when deciding whether to smile in an interaction. Therefore, the results do not allow us to draw conclusions with respect to how sincere or honest smiles might be perceived. However, it is likely to assume that such smiles will be more easily detected in the eyes area than our posed smiles. We believe that these results should be relevant for practitioners, entrepreneurs, and supervisors because they imply that wearing face masks will not lead to evaluations of sales representatives, negotiators, or teachers in a different way. Everybody can easily smile and detect smiles in the eyes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data, analysis code, and research materials are available at the OSF. <https://osf.io/hj3p5>

Acknowledgements

We gratefully acknowledge the research assistance provided by Ashley Durán, Silvia Ortiz, and Steffanny Romero. We thank Moritz Loewenfeld for his helpful comments. Financial support from the program “Inclusión productiva y social: programas y políticas para la promoción de una economía formal, código 60185, que conforma la Alianza EFI, bajo el Contrato de Recuperación Contingente No. FP44842-220-2018” is gratefully acknowledged. This research was further performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d’Avenir (ANR-11-IDEX-007) operated by the French National Research Agency.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.joep.2023.102660>.

References

- Andreoni, J., & Petrie, R. (2008). Beauty, gender and stereotypes: Evidence from laboratory experiments. *Journal of Economic Psychology*, 29(1), 73–93. <https://doi.org/10.1016/J.JOEP.2007.07.008>
- Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the United States of America*, 104(46), 17948–17953. <https://doi.org/10.1073/PNAS.0705435104>
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241–251.
- Bejarano, H., Gillet, J., & Rodríguez-Lara, I. (2021). Trust and trustworthiness after negative random shocks. *Journal of Economic Psychology*, 86, Article 102422. <https://doi.org/10.1016/J.JOEP.2021.102422>
- Biermann, M., Schulze, A., Unterseher, F., Atanasova, K., Watermann, P., Krause-Utz, A., ... Lis, S. (2021). Trustworthiness appraisals of faces wearing a surgical mask during the Covid-19 pandemic in Germany: An experimental study. *PLOS ONE*, 16(5), e0251393.
- Bonowski, T., & Minnameier, G. (2022). Morality and trust in impersonal relationships. *Journal of Economic Psychology*, 90, Article 102513. <https://doi.org/10.1016/J.JOEP.2022.102513>
- Burgess, A., & Horii, M. (2012). Risk, ritual and health responsabilisation: Japan’s “safety blanket” of surgical face mask-wearing. *Sociology of Health & Illness*, 34(8), 1184–1198. <https://doi.org/10.1111/j.1467-9566.2012.01466.x>
- Centorrino, S., Djemai, E., Hopfensitz, A., Milinski, M., & Seabright, P. (2015). Honest signaling in trust interactions: Smiles rated as genuine induce trust and signal higher earning opportunities. *Evolution and Human Behavior*, 36(1), 8–16. <https://doi.org/10.1016/J.EVOLHUMBEHAV.2014.08.001>
- Egan, M., Acharya, A., Sounderajah, V., Xu, Y., Mottershaw, A., Phillips, R., ... Darzi, A. (2021). Evaluating the effect of infographics on public recall, sentiment and willingness to use face masks during the COVID-19 pandemic: A randomised internet-based questionnaire study. *BMC Public Health*, 21(1), 1–10. <https://doi.org/10.1186/S12889-021-10356-0/FIGURES/3>
- Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, 24(6), 581–604.
- Ert, E., & Fleischer, A. (2020). What do Airbnb hosts reveal by posting photographs online and how does it affect their perceived trustworthiness? *Psychology & Marketing*, 37(5), 630–640. <https://doi.org/10.1002/MAR.21297>
- Feng, S., Shen, C., Xia, N., Song, W., Fan, M., & Cowling, B. J. (2020). Rational use of face masks in the COVID-19 pandemic. *The Lancet Respiratory Medicine*, 8(5), 434–436. [https://doi.org/10.1016/S2213-2600\(20\)30134-X](https://doi.org/10.1016/S2213-2600(20)30134-X)
- Filkuková, P., & Jørgensen, M. (2021). How to pose for a professional photo: The effect of three facial expressions on perception of competence of a software developer. *Australian Journal of Psychology*, 72(3), 257–266. <https://doi.org/10.1111/AJPY.12285>
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Forgas, J. P., & Laham, S. M. (2016). Halo effects. In *Cognitive Illusions* (2nd ed., pp. 286–300). Psychology Press.
- Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants’ attention to faces during the first year. *Cognition*, 110(2), 160–170. <https://doi.org/10.1016/j.cognition.2008.11.010>
- Giang, T., Bell, R., & Buchner, A. (2012). Does Facial Resemblance Enhance Cooperation? *PLoS ONE*, 7(10), 47809. <https://doi.org/10.1371/journal.pone.0047809>
- Godoy, R., Reyes-García, V., Huanca, T., Tanner, S., Leonard, W. R., McDade, T., & Vadez, V. (2005). Do smiles have a face value? Panel evidence from Amazonian Indians. *Journal of Economic Psychology*, 26(4), 469–490. <https://doi.org/10.1016/J.JOEP.2004.10.004>
- Golle, J., Mast, F. W., & Lobmaier, J. S. (2013). Something to smile about: The interrelationship between attractiveness and emotional expression. *Cognition & Emotion*, 28(2), 298–310. <https://doi.org/10.1080/02699931.2013.817383>
- Jaeger, B., Slegers, W. W. A., Evans, A. M., Stel, M., & van Beest, I. (2019). The effects of facial attractiveness and trustworthiness in online peer-to-peer markets. *Journal of Economic Psychology*, 75, Article 102125. <https://doi.org/10.1016/J.JOEP.2018.11.004>
- Krumhuber, E., Manstead, A. S. R., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial Dynamics as Indicators of Trustworthiness and Cooperative Behavior. *Emotion*, 7(4), 730–735. <https://doi.org/10.1037/1528-3542.7.4.730>
- Lavan, N., Mileva, M., & Mcgettigan, C. (2021). How does familiarity with a voice affect trait judgements? *British Journal of Psychology*, 112(1), 282–300. <https://doi.org/10.1111/bjop.12454>
- Marini, M., Ansani, A., Paglieri, F., Caruana, F., & Viola, M. (2021). The impact of facemasks on emotion recognition, trust attribution and re-identification. *Scientific Reports*, 11(1), 1–14. <https://doi.org/10.1038/s41598-021-84806-5>

- Marler, H., & Ditton, A. (2021). "I'm smiling back at you": Exploring the impact of mask wearing on communication in healthcare. *International Journal of Language & Communication Disorders*, 56(1), 205–214. <https://doi.org/10.1111/1460-6984.12578>
- Min, H., & Hu, Y. (2022). Revisiting the effects of smile intensity on judgments of warmth and competence: The role of industry context. *International Journal of Hospitality Management*, 102, Article 103152. <https://doi.org/10.1016/j.ijhm.2022.103152>
- Mueser, K. T., Grau, B. W., Sussman, S., & Rosen, A. J. (1984). You're only as pretty as you feel: Facial expression as a determinant of physical attractiveness. *Journal of Personality and Social Psychology*, 46(2), 469–478.
- Nobrega, M., Opice, R., Lauletta, M. M., & Nobrega, C. A. (2020). How face masks can affect school performance. *International Journal of Pediatric Otorhinolaryngology*, 138, Article 110328. <https://doi.org/10.1016/J.IJPORL.2020.110328>
- O'Doherty, J. P., Winston, J. S., Critchley, H., Perrett, D. I., Burt, M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, 41(2), 147–155.
- Orn, J., Scharlemann, P. W., Eckel, C. C., Kacelnik, A., Wilson, R. K., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, 22(5), 617–640. www.elsevier.com/locate/joep.
- Pavlova, M. A., & Sokolov, A. A. (2022). Reading Covered Faces. *Cerebral Cortex*, 32(2), 249–265. <https://doi.org/10.1093/cercor/bhab311>
- Póvoa, A. C. S., Pech, W., Viacava, J. J. C., & Schwartz, M. T. (2020). Is the beauty premium accessible to all? An experimental analysis. *Journal of Economic Psychology*, 78, Article 102252. <https://doi.org/10.1016/J.JOEP.2020.102252>
- Rhodes, G. (2006). The Evolutionary Psychology of Facial Beauty. *Annual Review of Psychology*, 57(1), 199–226. <https://doi.org/10.1146/ANNUREV.PSYCH.57.102904.190208>
- Sato, W., & Yoshikawa, S. (2010). Detection of emotional facial expressions and anti-expressions. *Visual Cognition*, 18(3), 369–388.
- Schmidt, K., Levenstein, R., & Ambadar, Z. (2012). Intensity of smiling and attractiveness as facial signals of trustworthiness in women. *Perceptual and Motor Skills*, 114(3), 964–978. <https://doi.org/10.2466/07.09.21.PMS.114.3.964-978>
- Sofianos, A. (2022). Self-reported & revealed trust: Experimental evidence. *Journal of Economic Psychology*, 88, Article 102451. <https://doi.org/10.1016/J.JOEP.2021.102451>
- Spezio, M. L., Rangel, A., Alvarez, R. M., O'Doherty, J. P., Mattes, K., Todorov, A., ... Adolphs, R. (2008). A neural basis for the effect of candidate appearance on election outcomes. *Social Cognitive and Affective Neuroscience*, 3(4), 344–352. <https://doi.org/10.1093/SCAN/NSN040>
- StataCorp. (2021). *Stata Statistical Software: Release 17*.
- Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, 21(3), 349–354. <https://doi.org/10.1177/0956797610362647>
- Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, 1124(1), 208–224. <https://doi.org/10.1196/ANNALS.1440.012>
- Tsantani, M., Podgajec, V., Gray, K. L. H., & Cook, R. (2022). How does the presence of a surgical face mask impair the perceived intensity of facial emotions? *PLoS ONE*, 17(1), e0262344.
- Tso, R. V., & Cowling, B. J. (2020). Importance of Face Masks for COVID-19: A Call for Effective Public Education. *Clinical Infectious Diseases*, 71(16), 2195–2198. <https://doi.org/10.1093/CID/CIAA593>
- Van Leeuwen, B., Noussair, C. N., Offerman, T., Suetens, S., Van Veelen, M., & Van De Ven, J. (2017). Predictably Angry—Facial Cues Provide a Credible Signal of Destructive Behavior. *Management Science*, 64(7), 3352–3364. <https://doi.org/10.1287/MNSC.2017.2727>
- Wang, Z., Mao, H., Li, Y. J., & Liu, F. (2017). Smile Big or Not? Effects of Smile Intensity on Perceptions of Warmth and Competence. *Journal of Consumer Research*, 43(5), 787–805. <https://doi.org/10.1093/JCR/UCW062>
- Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology*, 43(2), 303–311. <https://doi.org/10.1016/J.JESP.2006.02.004>