Routledge
Taylor & Francis Group

Basic and Applied Social Psychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hbas20

Face Coverings Differentially Alter Valence Judgments of Emotional Expressions

Nicholas R. Harp, Andrew T. Langbehn, Jeff T. Larsen, Paula M. Niedenthal & Maital Neta

BASIC AND APPLIED SOCIAL PSYCHOLOGY
https://doi.org/10.1080/01973533.2023.2221360

Face Coverings Differentially Alter Valence Judgments of Emotional Expressions

Nicholas R. Harp[a], Andrew T. Langbehn[b], Jeff T. Larsen[b], Paula M. Niedenthal[c], and Maital Neta[d]

[a]Yale University; [b]University of Tennessee-Knoxville; [c]University of Wisconsin-Madison; [d]University of Nebraska-Lincoln

ABSTRACT
Face masks that prevent disease transmission obscure facial expressions, impairing nonverbal communication. We assessed the impact of lower (masks) and upper (sunglasses) face coverings on emotional valence judgments of clearly valenced (fearful, happy) and ambiguously valenced (surprised) expressions, the latter of which have both positive and negative meanings. Masks, but not sunglasses, impaired judgments of clearly valenced expressions compared to faces without coverings. Drift diffusion models revealed that lower, but not upper, face coverings slowed evidence accumulation and affected differences in non-judgment processes (i.e., stimulus encoding, response execution time) for all expressions. Our results confirm mask-interference effects in nonverbal communication. The findings have implications for nonverbal and intergroup communication, and we propose guidance for implementing strategies to overcome mask-related interference.

The COVID-19 pandemic brought many changes to daily life, including the implementation of stay-at-home orders, social distancing guidelines, and the use of face masks to protect against the spread of disease. Although the days of government lockdowns and stay-at-home orders have largely passed, one change in daily behavior that could persist for some time is the use of face coverings to reduce viral transmission via respiratory droplets. Despite face coverings being highly effective as a preventative measure (for a review, see Howard et al., 2021), there may be unintended downsides to the use of such coverings. Most notably, face coverings occlude visual information, impacting communication through both verbal and nonverbal routes. For instance, masks reduce speech intelligibility (Caniato et al., 2021) and hamper emotion recognition (Carbon, 2020; Carbon & Serrano, 2021; Gori et al., 2021; Kim et al., 2022; Langbehn et al., 2022; Noyes et al., 2021; Roberson et al., 2012).

Indeed, previous research found that face coverings impact viewers' ability to judge emotional facial expressions accurately and effectively. Objects that obscure the face—ranging from a baby's pacifier to burqas or scarves—are known to hinder accurate judgments of emotional expressions (Fischer et al., 2012; Rychlowska et al., 2014). More recently, the coverings worn during the COVID-19 pandemic have been linked to a compromised ability to accurately judge emotion across several different emotional facial expressions (Carbon, 2020; Carbon & Serrano, 2021; Gori et al., 2021; Grundmann et al., 2021; Langbehn et al., 2022), an effect corroborated by lower confidence ratings of the judgments (Carbon, 2020). For instance, Langbehn et al. (2022) recently found that masks impacted judgments for expressions with facial activity predominantly in the lower half of the face, such as happiness and disgust, more severely than other expressions, such as anger and surprise.

Clearly, there are important effects of face coverings on nonverbal communication. However, it is worth noting some limitations of the research to date and opportunities to develop novel insights. Namely, previous research has (1) relied on expression recognition paradigms rather than affective (e.g., valence) judgments, (2) largely failed to compare ecologically valid upper vs. lower face coverings, and (3) focused on the outcome (e.g., emotion classification) rather than the judgment process underlying face covering effects.

Expression recognition and valence judgments

Prior work examining the impact of face masks primarily reports the effects of face coverings on the

CONTACT Nicholas R. Harp nicholas.harp@yale.edu Psychiatry, Yale University, New Haven, CT 06510, USA.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/01973533.2023.2221360.

accuracy of nonverbal communication, assessed via expression recognition of discrete emotion categories (e.g., Carbon, 2020; Carbon & Serrano, 2021; Gori et al., 2021; Kim et al., 2022; Langbehn et al., 2022; Roberson et al., 2012; Ruba & Pollak, 2020). Such an approach relies heavily on the idea that categorical emotions are conveyed through canonical expressions; and, while useful, leveraging affective judgments—rather than assessing category recognition directly—can lend additional insight into the accuracy of interpersonal judgments. For example, valence judgments non-only side-step the question about canonical expressions (Barrett et al., 2019), but they provide critical information about individuals' post-judgment intentions (e.g., approach-avoidance; Krieglmeyer et al., 2010). As such, measuring valence judgments likely better reflects early-stage emotion processing outside of the laboratory than category recognition tasks. That is, it is more important to first detect negative/aversive signals than it is to correctly categorize or label another's expression of negative emotion. Of course, it is worth noting that, even though valence (positive-negative) and motivational (approach-avoid) dimensions do not correspond perfectly (e.g., fearful and angry faces are both negatively valenced, but convey conflicting motivational signals; Carver & Harmon-Jones, 2009; Harmon-Jones & Allen, 1998), it is generally adaptive to approach positive and avoid negative stimuli.

In addition to lending insight into the accuracy of interpersonal judgments, measuring valence (as opposed to recognition of discrete emotion categories) can reveal notable individual differences. For example, valence judgments for expressions that convey clear positive (e.g., happy expressions) and clear negative (e.g., fearful expressions) signals are consistently positive and negative, respectively, but individuals differ in judgments of more emotionally ambiguous expressions (Neta et al., 2009). In other words, in the absence of clarifying context, some individuals tend to judge surprised facial expressions as negative (e.g., nearly stepping on a snake or spider), whereas others tend toward more positive judgments (e.g., finding extra cash in one's wallet). The individual differences in one's tendency to judge ambiguous information as more negative or positive constitute one's valence bias (Neta et al., 2009), which is temporally stable over periods of weeks to years (Harp et al., 2022; Neta et al., 2009).

As an individual difference measure, the valence bias—and valence judgments, more generally—might also be more sensitive to moderating variables than expression recognition tasks. For example, individuals endorsing higher levels of personality traits linked to negative affectivity (i.e., neuroticism) and lower levels of traits linked to social connectedness (i.e., extraversion) tend to show a more negative valence bias (Brock et al., 2022; Neta & Brock, 2021). As such, the effects of face coverings may vary as a function of these traits.

In addition to influences of personality, attitudes and beliefs—such as one's political orientation—might also moderate the impact of face coverings on valence bias. Indeed, judgments of ambiguous (surprised) expressions are more negative when participants are told that the faces belong to members of the opposing political party, at least in the U.S. (Basyouni et al., 2022). Because masks are likely to activate sociopolitical stereotypes/schema (e.g., Republicans compared to Democrats have more negative attitudes toward masks; Gelfand et al., 2022), valence judgments may vary along ideological lines—despite the absence of any effect on expression recognition (Langbehn et al., 2022).

Another putative moderator is a socioecological factor termed ancestral diversity—representing long-term human migratory patterns—which is associated with the degree to which geographic groups (e.g., states) both display and rely on facial signals to convey emotion signals (Niedenthal et al., 2023; Rychlowska et al., 2015). Specifically, individuals from more ancestrally diverse areas (e.g., New York) are likely to display and rely on the expressiveness of clear emotion signals to a greater degree than individuals from areas of relatively lower ancestral diversity (e.g., Tennessee; Niedenthal et al., 2018). As such, the effects of face coverings may be exacerbated for individuals from regions with higher ancestral diversity and mitigated for individuals from regions with lower ancestral diversity. Altogether, measuring valence judgments informs motivational tendencies, provides insight into individual differences in valence bias, and enhances sensitivity for assessing putative moderators in the impact of face masks on nonverbal communication.

Upper vs. lower coverings and ecological validity

A second limitation of prior research is that much of it has either not included an upper face covering condition (see Carbon, 2020; Carbon & Serrano, 2021; Gori et al., 2021; Grundmann et al., 2021; Langbehn et al., 2022) or they have used less ecologically valid coverings (e.g., randomly allocated windows or "bubbles"; Schyns et al., 2002). Including an upper face covering condition is necessary as both the upper and lower face contain dynamic musculature important for understanding facial displays of emotion (Barrett et al., 2019). Thus, without an upper face covering condition, the effect of lower coverings (masks) may be a result of occluding any part of the face, rather than specifically occluding the features of the lower part of the face.

Although Schyns et al. (2002) and Gibson et al. (2005) used random "bubbles" to examine the spatial location of the necessary features for discriminating facial characteristics (e.g., identity, expression) in both humans and animals, the approach suffers from low ecological validity. That is, the advantage of their approach are its high levels of experimental control, which comes at the cost of using the kinds of face coverings typical outside of the laboratory (e.g., face masks). Certainly, such artificially constrained views of faces rarely, if ever, occur in daily life. Nonetheless, this important foundational research shows that the features that are occluded—either at the top or bottom of the face—are important drivers of the effects on the perception of emotion or other facial characteristics (e.g., identity; Schyns et al., 2002). Making use of more ecologically valid face coverings (e.g., face masks, sunglasses) offers the opportunity to examine emotion judgments in a manner more closely resembling life outside of the laboratory (i.e., in the same setting in which our results would be applied; Trafimow & Osman, 2022). Thus, one opportunity to build upon earlier findings is to directly compare the effects of ecologically valid face coverings in the upper (sunglasses) versus lower (masks) halves of the face.[1]

Mechanisms of the judgment process

A third limitation of prior research is that it has not characterized the mechanisms *underlying* the impact of face coverings on emotion judgments, by, for example, analyzing response times (RTs). Although a few studies revealed a tendency for masks to slow RTs (e.g., Ziccardi et al., 2021), none have explored how the RT effect manifests in the judgment process. Sequential-sampling models, like the drift-diffusion model (DDM), allow the estimation of parameters that capture various components of the judgment-making process. Specifically, the DDM includes parameters that model the amount of evidence required to reach a judgment (i.e., threshold), the rate of evidence accumulation process (i.e., drift rate), *a priori* response/judgment biases (i.e., relative starting point), as well as the duration of non-judgment processes like stimulus encoding/sensory input and response execution/motor output time (Voss et al., 2004, 2010, 2015; Voss & Voss, 2007). As such, DDM allows researchers to determine whether—and how!—experimental manipulations affect distinct judgment processes. In other words, the DDM provides insight into the mechanisms by which experimental manipulations, such as the characteristics of various emotional expressions (e.g., clear vs. ambiguous valence) and covering types (e.g., masks vs. sunglasses), and/or person-level variables (e.g., personality traits, political orientation), impact the judgment process. Thus, the DDM provides a more informative analysis than examining judgments and/or RT in isolation (see Johnson et al., 2017 for a primer on DDM in social and personality psychological contexts and a more thorough description of DDM in the Method section).

The present study

Here, we investigate the influence of different face coverings (masks, sunglasses, no coverings) on valence judgments of facial expressions with either clear positive (happy), clear negative (fearful), or ambiguous (surprised) meaning. Although we did not preregister specific predictions about the effects of face coverings on judgments of happy and fearful expressions (which were included as clearly valenced anchors), we expected that face masks would lead to less accurate valence judgments for both of these expressions (see Langbehn et al., 2022; Roberson et al., 2012). This is especially the case for happy expressions, which contain most of their activity in the lower half of the face (Langbehn et al., 2022).

We did, however, preregister the hypothesis that masks (i.e., occluding the bottom half of the face), but not sunglasses (i.e., occluding the top half of the face), will be associated with more negative judgments of surprised facial expressions. This prediction stems from evidence that visual search patterns (i.e., fixating on the mouth vs. the eyes) play a causal role in determining judgments of surprised expressions. Specifically, individuals who make faster fixations to the mouth—a feature that discriminates surprise from fear—tend to judge surprised expressions more

[1]Some researchers have, indeed, compared the effects of sunglasses and masks on emotion recognition, but the previously reported effects are pooled across emotional expressions (Roberson et al., 2012) or rely on a basic emotion approach (i.e, categorical response; Kim et al., 2022; Ruba & Pollak, 2020), limiting inferences about expression-specific effects (but see Noyes et al., 2021, which used a "match" vs. "no match" judgment approach and examined recognition by expression).

positively (Neta et al., 2017). Indeed, fearful and surprised expressions share morphological similarities in the upper half of the face (e.g., widening of the eyes), as evidenced by overlapping facial AUs (i.e., inner/-outer brow and upper eyelid raising, AUs 1, 2, and 5; Du & Martinez, 2015), and action in the lower half of the face is helpful for distinguishing between the expressions (Farah et al., 1998). Thus, fixation patterns play a causal role in determining judgments of surprised expressions, meaning that the face coverings worn during the COVID-19 pandemic may make people more likely to judge surprised expressions as negative. We used DDM to further examine *how* the judgment process varied among experimental conditions. In exploratory analyses, we assessed whether personality traits, political orientation, or ancestral diversity moderated mask-related effects (see Supplemental Material section 2).

Method

Participants

Participants were recruited from Amazon's Mechanical Turk (Mturk) and invited to participate in an eligibility screener (US$0.10), with an opportunity for eligible participants to earn a bonus (US$1.90). Only workers aged 18 years or older who were free of previous psychological or neurological diagnoses were invited to participate in the bonus. 577 workers completed the screener, 222 were ineligible and 150 were rejected for failing to complete the study within 1.5 hours, leaving 205 final participants (pre-registered target for recruitment: 200). Five participants were removed prior to analysis due to having completed too few trials, and an additional 54 were removed due to inaccurate judgments of the clear valence stimuli (both described below). Demographic characteristics of the final sample ($n = 146$; pre-registered target for analysis (i.e., after data quality check): 150) are shown in Table 1. All participants provided informed consent in accordance with the Declaration of Helsinki and all research procedures were approved through the University of Nebraska-Lincoln's Institutional Review Board (Approval #20150114791EP).

Table 1. Demographic characteristics.

Demographic	Final sample ($n = 146$)
Age	M(SD) = 40.87(12.30); range = 18-77 years
Gender	Female (76), Male (69), Prefer not to answer (1)
Race	American Indian or Alaskan Native (1), Asian (19), Black (7), Other (3), White (114), Prefer not to answer (2)
Ethnicity	Latino/Hispanic (6), Not Latino/Hispanic (139), Prefer not to answer (1)

Stimuli and measures

One hundred eighty facial expressions from 111 unique identities were selected from the NimStim (25 identities, 44 of 180 expressions; Tottenham et al., 2009), Karolinska Directed Emotional Faces (26 identities, 28 of 180 expressions; Lundqvist et al., 1998), Radboud Faces Database (15 identities, 19 out of 180 expressions; Langner et al., 2010), and Umea (37 identities, 68 out of 180 expressions; Samuelsson et al., 2012) sets, as well as still images taken from a set of dynamic emotional expressions described elsewhere (8 identities, 21 out of 180 expressions; Langbehn et al., 2022). All faces displayed either fearful, happy, or surprised expressions. Face masks and sunglasses were added to each expression in Adobe Photoshop. Each stimulus appeared only once; that is, any individual participant only saw a given expression in a single face covering condition (counterbalanced across participants). See Figure 1 for sample images from each condition.

Procedure

The task and surveys were hosted on Gorilla Experiment Builder (Anwyl-Irvine et al., 2020), and

Figure 1. Sample stimuli from face covering and emotional expression conditions. Happy, fearful, and surprised expressions (top to bottom) without coverings, with masks, or with sunglasses (left to right). NimStim model 01 F is shown here (Tottenham et al., 2009).

accessible to participants using a computer (no phones or tablets). Participants first provided informed consent, followed by a brief (4 trial) practice task and a (180 trial) main task. Both the practice and main tasks used a two-alternative forced choice design such that participants judged facial expression images as either positive or negative via keyboard button press. Participants were randomly assigned to one of two counterbalanced orders for response options (L = Positive & A = Negative; L = Negative & A = Positive). There were 60 fearful, 60 happy, and 60 surprised expressions presented to each subject. Within each expression, equal subsets of trials were shown with a surgical mask, sunglasses, or no face coverings (20 trials each). Each image was presented for 500 ms and preceded by a 1,500 ms fixation cross, consistent with extensive prior work (e.g., Harp et al., 2021; Neta et al., 2013). Failure to respond within 2000 ms resulted in the task advancing to the next trial without recording a response (response omission), and only the first response was retained for analysis in the event of multiple responses on a single trial. The order of the stimuli was pseudorandomized, such that there was an even number of fearful or happy trials preceding surprise trials in all versions to rule out order or context effects (e.g., Neta et al., 2011). After the task, participants completed the surveys described in the Supplemental Material.

Drift diffusion modeling

The drift-diffusion model (DDM) offers an advantageous framework for decomposing response and response time (RT) data into interpretable parameters. Specifically, it decomposes responses and RTs from tasks with binary judgments into parameters estimated in a quantitative model fit to each individuals' data (Lerche & Voss, 2018; Ratcliff, 1978; Voss et al., 2015). Because the model fits to individual-level data, the parameter estimates can be averaged across persons and compared between experimental conditions to inform conclusions about how experimental manipulations impact the judicial process; inferences that cannot be made from responses or RTs alone (Johnson et al., 2017; Ratcliff, 1978; Voss et al., 2004, 2010, 2015; Voss & Voss, 2007). For example, the DDM models correct and incorrect responses, the latter of which tend to be removed prior to traditional RT analyses (Johnson et al., 2017). Additionally, the parameters estimated via the DDM provide a more objective measure of cognitive processes than self-report; that is, the DDM is not susceptible to participant biases/motivations (e.g., demand characteristics; Johnson et al., 2017).

There are four primary parameters in the DDM, and each parameter represents a different possible source of variability in the judgment process (Lerche & Voss, 2018). First is the threshold (a), which represents the distance between the lower (0; here, "negative") and upper (a; here, "positive") judgment boundaries. In short, the threshold represents the amount of evidence required to reach a particular judgment in a binary judgment task. Thresholds can vary across persons and experimental conditions. Between-person variability arises due to individual differences in cautiousness, or conservativeness, of judgment-making. Specifically, individuals who emphasize accuracy over speed will show a higher threshold (e.g., older adults often show a higher threshold than younger adults; Ratcliff et al., 2001; Spaniol et al., 2006). On the other hand, thresholds may differ within persons due to experimental manipulations (e.g., instructions to emphasize accuracy over speed would result in a higher, more conservative threshold a; Ratcliff & Rouder, 2000; Wagenmakers et al., 2008). Here, we held the threshold constant across experimental conditions, such that the DDM estimates only a single threshold per participant, given that participants were instructed to respond quickly *and* accurately for all trials.

Whereas threshold (a) represents the conservativeness of one's judgments, the drift rate (v) represents the speed at which evidence accumulates toward one of the judgment boundaries. There are two characteristics of drift rate that inform the judgment process: (1) the sign and (2) the magnitude. Given that there are both upper and lower judgment boundaries, evidence can accumulate in either the upper (positive-signed drift rate) or lower (negative-signed drift rate) direction. Thus, the sign of the drift indicates whether evidence tends to accumulate toward the upper (positive sign) or lower (negative sign) boundary. The magnitude (i.e., absolute value) of the drift rate indicates the speed of evidence accumulation, where a larger magnitude corresponds with faster accumulation and a smaller magnitude corresponds with slower accumulation. Experimental manipulations common to both cognitive and social psychology paradigms influence drift rates. For example, experimental manipulations which increase difficulty (e.g., judging color dominance in a visual sample with 51.5% [more difficult] vs. 53% dominance [less difficult]) slow the rate of evidence accumulation (i.e., reducing drift rate; Voss et al., 2004). Similarly, judgments that activate stereotypic associations (e.g., between people of color and violence) impact drift rates in a first-person shooter task (i.e., drift rates are faster for "gun" vs. "non-gun" judgments on non-gun trials with Black compared to White targets; Pleskac et al., 2018). Taken together, we expected to observe slower drift rates for masked faces because, in the case of valence judgments of static faces, obscuring the shape of the mouth, eyebrows, and so forth should hinder stimulus discrimination (and perhaps activate social schemas depending upon one's political orientation; Gelfand et al., 2022).

Third, the DDM models a relative starting point (z_r). The relative starting point captures whether participants show a judgment bias, such that one judgment requires less evidence. For instance, there might be differences in relative starting points when one judgment is linked to rewards or losses (e.g., Voss et al., 2008). Differences in relative starting point are thought to reflect *a priori* judgment biases rather than the perceptual process captured by drift rate (v). Because our experimental design was balanced and pseudorandomized, such that participants viewed an equal number of positive, negative, and ambiguous expressions, we fixed starting point to the midpoint (i.e., assuming no *a priori* judgment bias) rather than estimating it per condition.

Last, differences in non-judgment processes (d) are modeled. The non-judgment processes represent processes that occur either before (i.e., stimulus encoding/sensory input) or after (i.e., response execution/motor output) the evidence accumulation process that produces a judgment (Lerche et al., 2017; Theisen et al., 2021). As an example, consider again the differences in performance of younger and older adults on basic cognitive (e.g., memory) tasks. Above, we discussed a finding that response cautiousness (i.e., threshold, a) tends to be greater for individuals in older adulthood. Similarly, there is a well-replicated finding that the duration of non-judgment processes tends to be greater in older than younger adults (for a meta-analysis, see Theisen et al., 2021). The DDM has helped to tease apart whether the slower RTs of older adults on cognitive tasks are attributable to deficits in judgment-related processes (threshold, a or drift, v), non-judgment processes (e.g., impaired vision, reduced motor capabilities), or some combination of both. Because non-judgment processes, by definition, exclude the evidence accumulation process, the parameter is sensitive to differences in RT but not accuracy. Thus, the non-judgment parameters can vary without impacting the judgment—or evidence accumulation—process itself. Because masks occlude part of the face that would otherwise be encoded during stimulus encoding/sensory input, we estimated differences in the speed of response execution for each experimental condition (d) in the DDM.

To implement DDM, we used fast-dm (Voss & Voss, 2007) and estimated parameters from our trial-wise data, focusing on drift rate (v). Parameters were estimated using the Kolmogorov-Smirnov method (Lerche et al., 2017; Voss et al., 2004, 2015). As in previous work (Voss et al., 2004), and in line with current guidance in implementing fast-dm models (Voss et al., 2015), model fit was assessed for each participant and three participants with a significant Kolmogorov-Smirnov test result ($p < .05$) were removed due to bad model fit. We selected the Kolmogorov-Smirnov test given our moderate number of trials (i.e., 180 trials), as maximum likelihood optimization is particularly sensitive to outlier RTs and chi-square optimization requires substantially larger numbers of trials (e.g., 1000+ trials; Voss et al., 2015). Although we opted for a trial number closer to the lower boundary needed for precise estimation given our modeled parameters (i.e., threshold, drift, and non-judgment processes) and optimization approach (i.e., Kolmogorov-Smirnov; Lerche et al., 2017), it is possible that some participants experienced fatigue or boredom during the task. Such fatigue or boredom may have contributed to the notable rate of exclusions described above. One additional participant was not analyzed because fast-dm could not generate parameter estimates from the data (i.e., there was no data for some conditions). The resulting parameter estimates from retained participants ($n = 142$) were analyzed as described in the following section.

Data analysis

Preregistration (https://osf.io/re82d) and deidentified data with analysis scripts (https://osf.io/2wn7r/) are available on the Open Science Framework. All data cleaning, analyses, and visualizations were completed in R (version 4.1.1; R Core Team, 2022) and fast-dm (Voss & Voss, 2007). As in previous work and as specified in our preregistration, we removed trials with RT either less than 250 milliseconds or greater than three standard deviations from the participant's mean RT and then calculated the percentage of faces in each category (e.g., masked surprised faces) that were judged to be positive or negative. Five participants retained less than 75% of the total number of trials and were thus removed from the analyses. Next, we assessed the accuracy of responses to the clearly valenced stimuli (unmasked fearful and happy faces), removing 54 participants that failed to judge these images with greater than 60% accuracy, as in previous work (Harp et al., 2021; Neta et al., 2019). Notably, the percentage of excluded participants (27%; 54 of 200) was close to our expected rate of exclusion in our pre-registration (25%; 50 of 200). Additionally, others have reported somewhere between 10-50% exclusion rates in online samples (Curran, 2016); thus, our manipulation/data quality check appears to have succeeded.

We took a primarily descriptive approach to data analysis. Given concerns about the null hypothesis significance testing framework (see Trafimow & Marks, 2015), we report descriptive statistics and effect sizes (Cohen's d) rather than frequentist inferential values. We analyzed both the percent positive judgments and the DDM parameters (i.e., drift [v] and differences in non-judgment processes [d]) data by computing descriptive statistics (i.e., means, standard deviation, and skew) for the within-subject effects of expression (fearful, happy, surprised) and covering (mask, none, sunglasses). Plots were completed using ggplot2 (Wickham, 2009). We also describe a more traditional RT analysis and several exploratory moderation analyses, testing for effects of personality, political orientation, and ancestral diversity on percent positive judgments in the Supplemental Material (section 2).

Results

Valence judgments

Histograms of valence judgment data are in Figure 2. The bar plot for the percentage of faces judged as positive versus negative is shown in Figure 3. Judgments for surprised faces (M(SD) = 79.37(19.38), skew = −1.41) were less negative than judgments for fearful faces (M(SD) = 90.01(12.98), skew = −2.05; $d = 0.59$) and more negative than judgments for happy faces (M(SD) = 18.49(20.16), skew = 2.08; $d = -3.08$). The effect sizes for the main effects of expression were much larger than the effect sizes for the main effects of covering, which were mostly negligible. Compared to faces without coverings (M(SD) = 61.93(37.71), skew = −0.53), masks led to slightly more negative judgments (M(SD) = 65.27(33.90), skew = −0.63; $d = 0.09$) whereas sunglasses seemingly had no impact (M(SD) = 60.67(36.77), skew = −0.50; $d = 0.03$). Masks also led to slightly more negative judgments compared to sunglasses ($d = 0.13$).

Next, we examined whether there were interactive effects of expression and covering. Descriptive statistics for each cell of the experimental design (expression x covering) are available in Table 2. As shown in Figure 3, our preregistered hypothesis that masked surprised faces is more likely to be judged as negative than surprised faces without masks was not observed. Nonetheless, there were apparent effects of face covering for both fearful and happy faces. Masks led to more inaccurate (i.e., positive) judgments of fearful faces, such that the percentage of negative judgments was greater for no coverings ($d = -0.39$) and for sunglasses ($d = -0.26$). Sunglasses did not appear to impact judgments compared to no coverings ($d = 0.16$). Masks also led to more inaccurate (i.e., negative) judgments of happy faces compared to no coverings ($d = 0.75$) and sunglasses ($d = 0.71$). Again, sunglasses did not appear to impact judgments compared to no coverings ($d = -0.12$). Unexpectedly, masks did not affect judgments of surprised faces ($d = -0.12$). Sunglasses, however, led to somewhat less negative judgments of the surprised expressions ($d = 0.19$) compared to no coverings, though the effect was quite small. There was no meaningful difference between judgments of surprised expressions with masks versus sunglasses ($d = 0.05$).

Reaction times

Histograms of reaction times are shown in Figure 4 and statistical analyses are available in the Supplementary Material section 1. In brief, masks slowed RTs compared to sunglasses and faces without coverings, and this pattern was evident for each emotional expression. Thus, masks slowed valence judgments.

Drift rate (v)

To understand the mechanisms through which face masks impacted the judgment process, we submitted responses and RTs to DDM to assess the impacts of masks on drift rate (i.e., the rate of evidence accumulation toward a judgment boundary). The drift rate varied across the three expressions. Drift rates showed greater evidence accumulation toward the "negative" (lower) judgment boundary for fearful (M(SD) = −3.06(1.69), skew = −0.77) than surprised expressions (M(SD) = −1.95(1.53), skew = −0.02; $d = -0.68$), although evidence tended to accumulate toward the lower boundary (i.e., "negative" judgment) for both expressions. On the other hand, the average drift rate

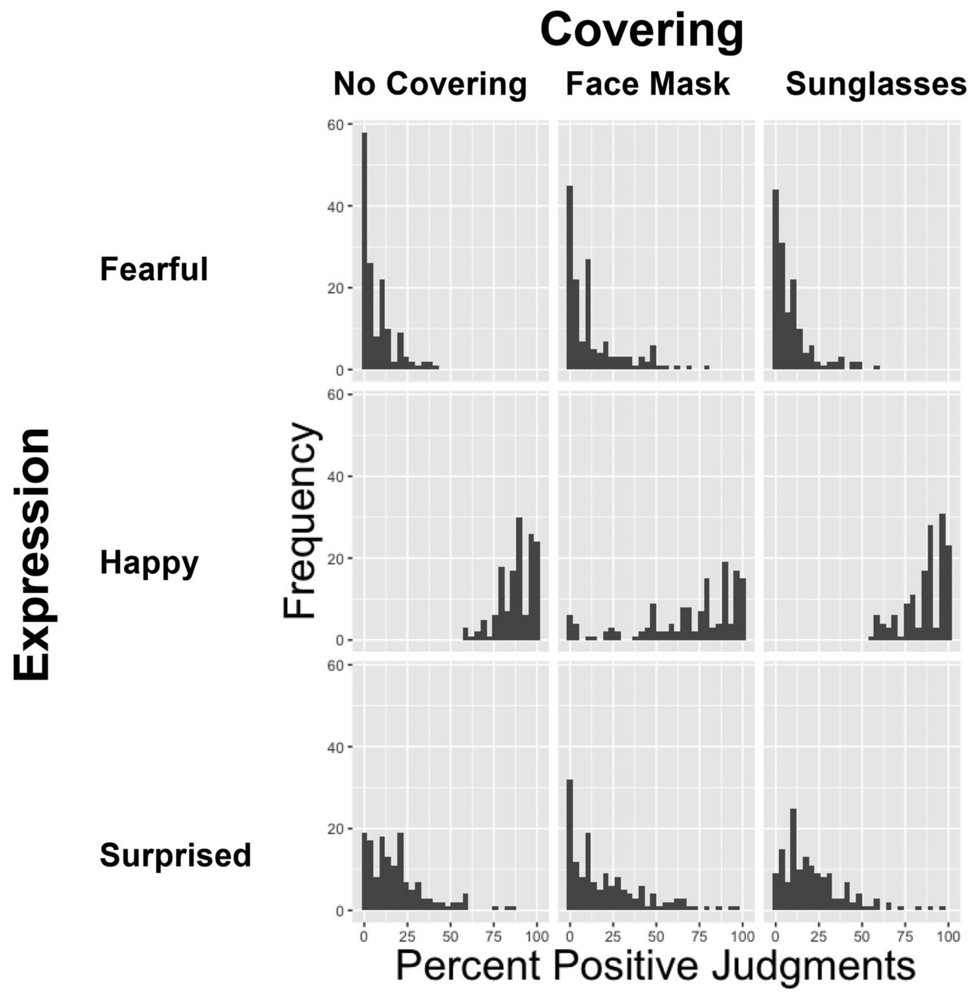

Figure 2. Histogram of the percentage of fearful, happy, and surprised faces with masks, no face coverings, and sunglasses that participants judged to be expressing positive emotions. Frequency values represent a number of participants.

for happy expressions was positive (M(SD) = 2.36(1.72), skew $= -0.75$), indicating that evidence tended to accumulate toward the upper boundary (i.e., "positive" judgment). The differences in drift rates for happy compared to fearful ($d = 3.18$) and surprised expressions ($d = 2.64$) were large.

Next, we examined the effects of face coverings for each expression (Figure 5). Descriptive statistics for each cell of the experimental design of expression and covering are available in Table 3. For fearful expressions, masks tended to slow drift rate compared to no coverings ($d = 0.41$) and sunglasses ($d = 0.23$). Sunglasses, however, did not slow drift rates relative to no coverings ($d = -0.17$). For the happy expressions, masks slowed drift rates compared to no coverings ($d = -0.77$) and sunglasses (M(SD) = 2.88(1.35), $d = -0.81$). Sunglasses did not impact the drift rate for the happy expressions ($d = -0.09$) compared to no coverings. For surprised expressions, masks tended to slow drift rate compared to no coverings ($d = 0.23$). There was no effect of sunglasses on surprised expressions compared to no coverings ($d = -0.12$) or masks ($d = 0.11$).

Non-judgment processes (d)

Differences in the duration of non-judgment processes (i.e., stimulus encoding/sensory input and response execution/motor output) are plotted in Figure 6. Descriptive statistics for each cell of the experimental design of expression and covering are available in Table 4. Note that plotted values indicate the *difference* in non-judgment time, such that positive values of *d* indicate that non-judgment processes are faster for judgments linked to the upper threshold (i.e., "positive" judgment) than for judgments linked to the lower threshold.

Masks significantly impacted differences in non-judgment time. Fearful and surprised expressions with either no coverings or sunglasses had positive-signed estimates, meaning that stimulus encoding and motor execution were faster for judgments linked to the

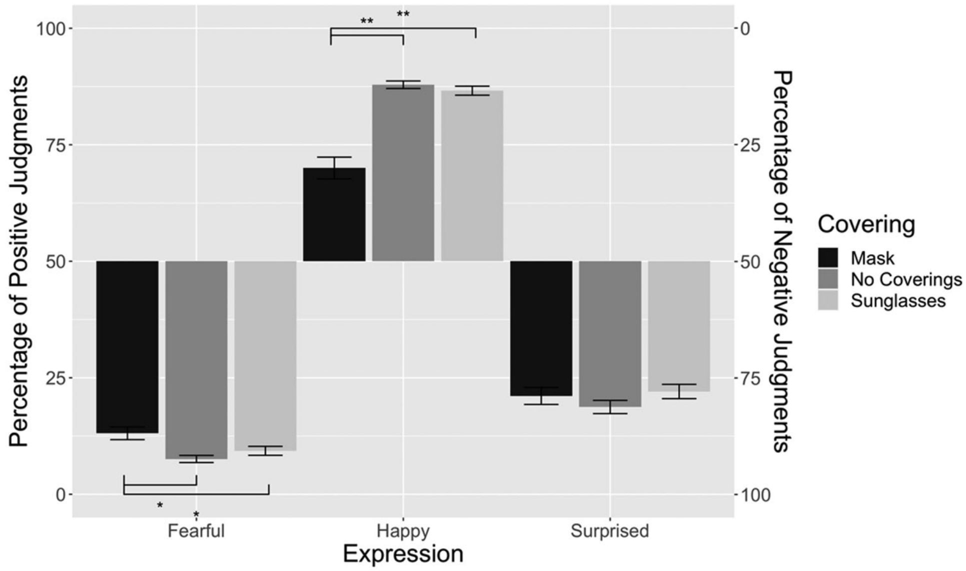

Figure 3. Percentage of positive judgments for each expression and condition. Bars above 50% indicate that >50% of faces were judged positive. Bars below 50% indicate that >50% of faces were judged to be negative. Masks impaired valence judgments for both fearful and happy relative to faces without coverings ($ds \geq 0.39$) and sunglasses ($d \geq 0.26$). Surprised faces without coverings were judged as somewhat more negative than surprised faces with sunglasses ($d = 0.19$). *** = large effect size, ** medium effect size, * small effect size.

Table 2. Descriptive statistics for percent negative judgments.

		Covering: M(SD) [skew]		
		No covering	Mask	Sunglasses
Expression	Fearful	92.42(9.21) [−1.45]	86.91(16.49) [−1.68]	90.68(11.63) [−1.99]
	Happy	12.11(9.77) [0.85]	29.98(28.16) [1.09]	13.39(11.66) [0.93]
	Surprised	81.27(17.19) [−1.52]	78.90(22.04) [−1.24]	77.94(18.55) [−1.47]

lower threshold (i.e., "negative" judgment) than the upper ("positive") threshold. Masks reversed the effect, such that both fearful ($ds \geq 0.30$) and surprised ($ds \geq 0.51$) expressions with masks had *negative*-signed estimates, meaning that stimulus encoding and motor execution were faster for judgments linked to the upper threshold (i.e., "positive" judgment) than the lower threshold. Happy expressions showed the inverse pattern; those with no coverings or sunglasses showed a negative-signed estimate (faster for "positive" judgments) but those with masks showed a positive-signed estimate (faster for "negative" judgments; $ds \geq 0.54$). The effect sizes for fearful expressions were small, whereas the effect sizes for surprised and happy expressions were medium. There were no meaningful effects of sunglasses on non-judgment processes for any of the expressions. In sum, masks resulted in faster erroneous judgments for the fearful and happy expressions, and faster "positive" judgments for the surprised expressions, compared to faces without coverings or with sunglasses.

Discussion

In this study, we assessed the impact of upper and lower face coverings on valence judgments of three different emotional expressions. We also submitted response and RT data to drift-diffusion modeling (DDM) to explore the underlying mechanisms of face-covering interference. Our pre-registered hypothesis that surprised faces are judged as more negative when face masks occluded the mouth—a distinguishing characteristic from fearful expressions—was not supported. Face masks did, however, lead to more positive judgments of fearful expressions and more negative judgments of happy expressions. Face masks also resulted in impairments in the evidence accumulation and non-judgment processes (e.g., stimulus encoding/sensory input). This pattern is consistent with prior research showing that masks impede communication of these expressions that convey relatively clear valence, and shed light on the mechanisms through which the impediment manifests. We discuss how masks impacted the judgment (i.e., slowed evidence accumulation) and non-judgment processes, with specific guidance on strategies that hold promise for overcoming mask-induced interference.

Face coverings impact valence judgments

Building upon recent evidence that face coverings impair people's ability to recognize discrete emotional facial expressions (Carbon, 2020, 2020; Fischer et al.,

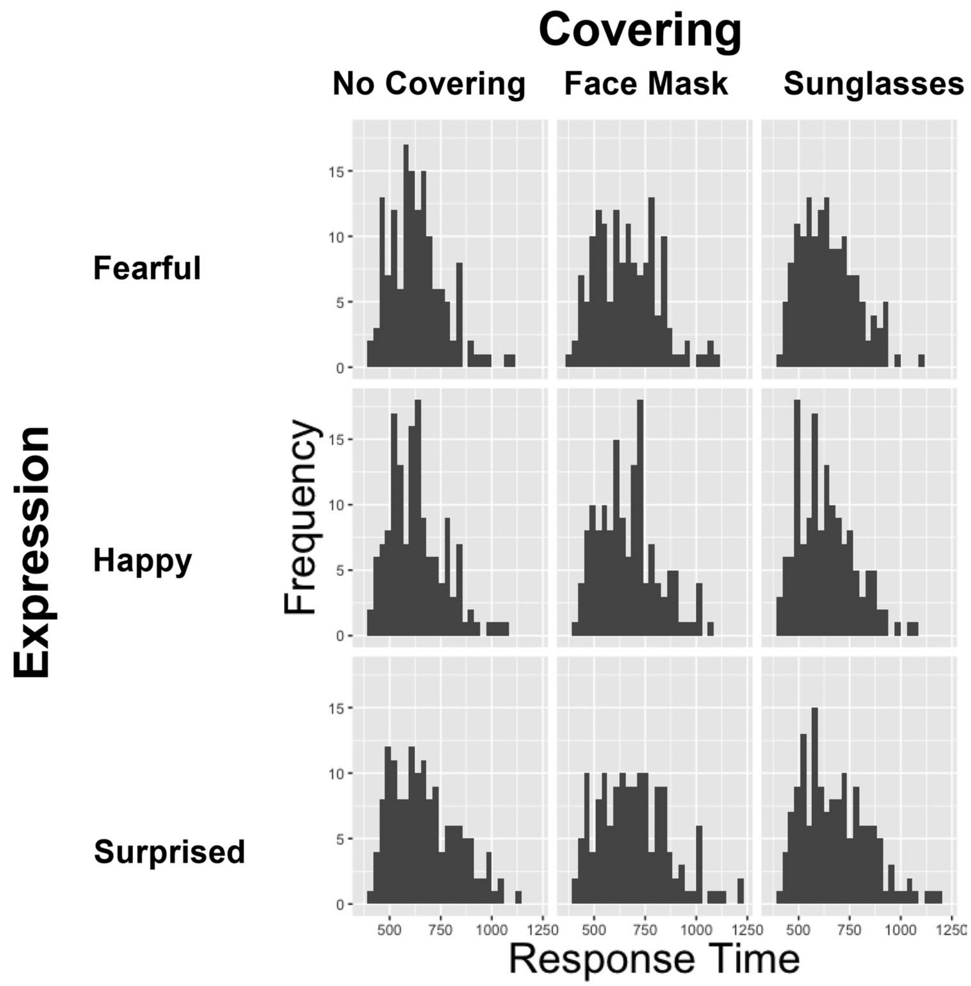

Figure 4. Histogram of mean response times for fearful, happy, and surprised faces with masks, no face coverings, and sunglasses. Frequency values represent the number of participants.

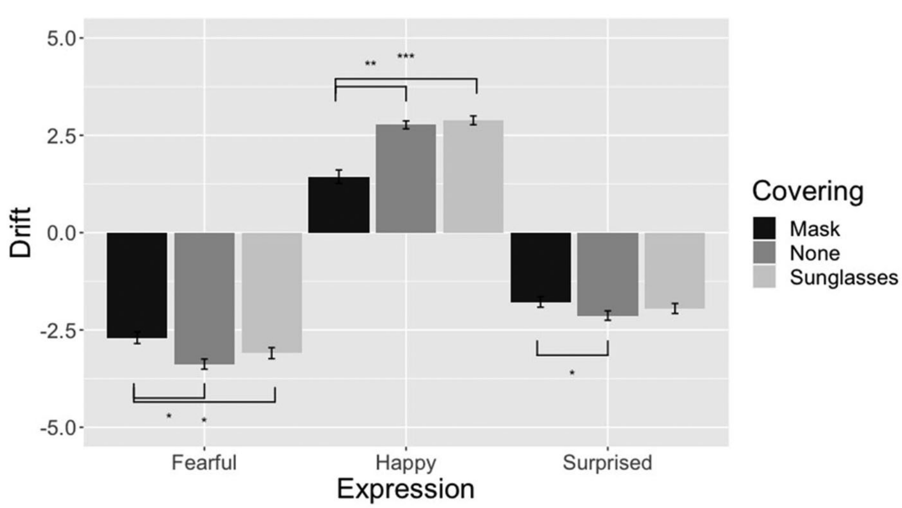

Figure 5. Drift rates for each expression and condition. Larger absolute values of drift indicate faster evidence accumulation toward either the upper (i.e., "positive" judgment) or lower boundary (i.e., "negative" judgment). Masks slowed drift rates relative to faces without coverings or with sunglasses. ***large effect size, **medium effect size, *small effect size.

2012; Gori et al., 2021; Grundmann et al., 2021; Kim et al., 2022; Langbehn et al., 2022; Noyes et al., 2021; Roberson et al., 2012; Ruba & Pollak, 2020; Rychlowska et al., 2014), our results provide a novel contribution by extending this work to show an impairment of basic valence judgments. Given our use of dimensional valence judgments rather than discrete emotion category recognition, we provide novel evidence that masks interfere not only with recognition accuracy but also with a more fundamental component of affective experience and judgments (Russell, 2003). As such, the findings have implications for how masks may impact interpersonal interactions via the relationship between valence judgments and motivational action tendencies (e.g., to approach-avoid). For instance, participants were more than twice as likely to judge a smiling face as negative when the target was wearing a mask as opposed to no covering or sunglasses (see Table 2), corresponding to a large effect size. Such misjudgments could have negative consequences for real-world social encounters. For instance, individuals may misinterpret a passerby's friendly smile as a threat or, conversely, fail to adequately interpret another's expression of fear and overlook a potential threat in the immediate environment. In other words, both misjudgments threaten adverse consequences (e.g., lost social opportunities, and increased vulnerability to threats). Relatedly, others report that masks influence judgments other than valence. For instance, masks lead people to judge expressions conveying negative valence as more trustworthy and likeable when masked than unmasked (Grundmann et al., 2021). Thus, face masks pose challenges to nonverbal communication, making faces that convey positive meaning look less positive, but also making faces that convey negative meaning look less negative.

Unexpectedly, face masks did not meaningfully impact valence judgments of surprised expressions. Given that faster fixations on the mouth are associated with more positive judgments of surprise (Neta et al., 2017; Neta & Dodd, 2018), we expected face masks would lead to more negative judgments of surprise

Table 3. Descriptive statistics for drift rates (*v*).

		Covering: M(SD) [skew]		
		No covering	Mask	Sunglasses
Expression	Fearful	−3.38(1.55) [−0.95]	−2.70(1.77) [−0.57]	−3.10(1.69) [−1.08]
	Happy	2.77(1.21) [0.18]	1.43(2.08) [−0.59]	2.88(1.35) [0.60]
	Surprised	−2.13(1.44) [0.07]	−1.78(1.62) [0.02]	−1.95(1.53) [−0.23]

Table 4. Descriptive statistics for differences in non-judgment processes (*d*).

		Covering: M(SD) [skew]		
		No covering	Mask	Sunglasses
Expression	Fearful	−0.01(0.07) [−0.17]	0.02(0.08) [0.06]	−0.01(0.08) [−0.68]
	Happy	0.03(0.09) [0.37]	−0.03(0.09) [0.18]	0.02(0.08) [0.30]
	Surprised	−0.02(0.08) [0.17]	0.03(0.08) [1.02]	−0.02(0.09) [−0.49]

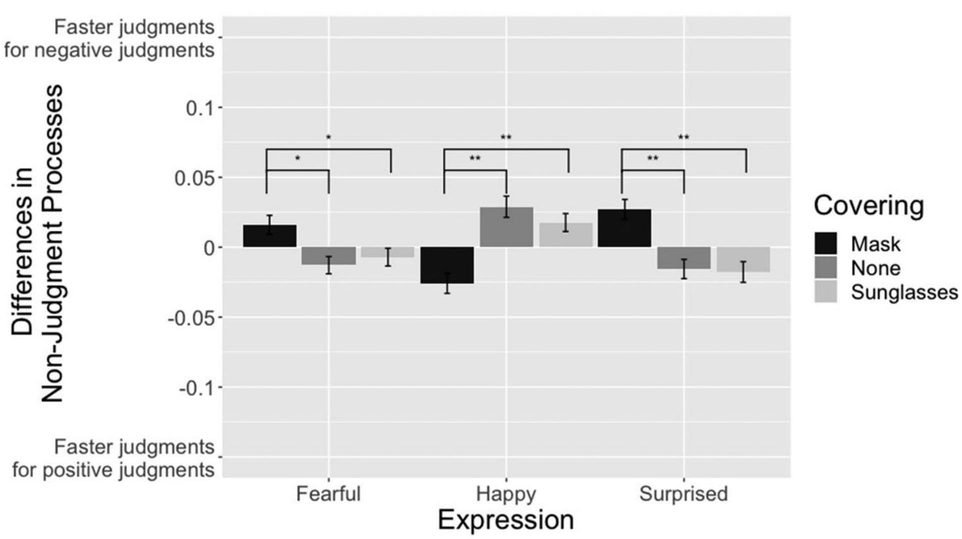

Figure 6. Differences in non-judgment processes (*d*; e.g., response execution) for each expression and condition. Positive values indicate that response execution is faster for judgments linked to the upper threshold (i.e., positive judgment) than for judgments linked to the lower threshold. Masks resulted in a reversal of the sign of the differences in non-judgment processes for all three expressions. Erroneous judgments for fearful and happy expressions, as well as "positive" judgments for surprised expressions, were faster than responses for the alternative judgment. ***large effect size, **medium effect size, *small effect size.

expressions. One possible explanation for our results is that morphological similarities between fearful and surprised expressions led participants to judge surprise as more negative regardless of face coverings, perhaps due to confusing the expressions with each other. Such a pattern resembles the exacerbated interference of masks for other perceptually similar faces (e.g., happiness and disgust; Barrett et al., 2019) seen in prior work (Langbehn et al., 2022). Indeed, the group average for the percentage of negative judgments for surprised expressions without face coverings surpassed 75% in these data, which is noticeably higher than the approximately 50% group average in other samples (Neta & Brock, 2021). Additionally, it could be that face masks create a situational context prompting more negative judgments of ambiguity. That is, even surprised faces without coverings might have been judged as more negative, given that masks on temporally proximal faces could suggest the possibility of disease exposure and create a threatening context (Neta et al., 2011). Lastly, occluding the mouth slowed RTs (see Supplemental Material section 1). The putative negativity we predicted for masked surprised expressions may have conflicted with the tendency for slower responses to be linked to more positive judgments of surprise (Neta & Tong, 2016); in other words, the effects may have canceled each other out.

Face masks impact both judgment and non-judgment processes

To characterize the mechanisms through which face coverings influence the judgment process, we fit DDMs to the response and RT data. Given that both masks and sunglasses occlude visual cues from the face, we expected faster evidence accumulation (i.e., larger magnitude drift rate) for expressions without coverings compared to expressions with coverings. However, sunglasses did not meaningfully impact drift rates relative to faces without coverings, perhaps because the distinguishing characteristic between surprised and fearful as well as the most notable characteristic of happy expressions—the mouth—remained visible in the sunglasses condition. It could also be that the greater impact of masks relative to sunglasses is due to the distribution of AUs in the lower (eighteen AUs) and upper face (twelve AUs; Ekman & Friesen, 1978), or that sunglasses are more familiar and thus had a smaller impact than masks.

Masks, on the other hand, impacted drift rates for each expression, though the degree of interference varied. Specifically, masks slowed evidence accumulation for the happy and fearful faces compared to those without coverings or with sunglasses (consistent with effects in valence judgments and RTs; see Supplemental Material section 1), whereas the effect of masks for surprised faces was smaller. One interpretation of the range in effect sizes on drift rates is that masks made stimulus discrimination more difficult for the clear expressions (Voss et al., 2004), but that the ambiguity inherent to surprised expressions superseded the impact of masks on evidence accumulation. In other words, surprised expressions—regardless of coverings—tended to have slower absolute drift relative to the clearly valenced expressions, perhaps resulting in a near-floor effect and lessening the impact of additional interference by masks or other coverings on evidence accumulation (at least in valence judgment tasks). Fortunately, there are methods for offsetting differences in evidence accumulation (e.g., instructions to emphasize accuracy and thus increasing the amount of evidence needed to reach a judgment threshold; Johnson et al., 2017), suggesting that the effects of masks might be overcome with sufficient experience or tailored interventions.

Additionally, we found that face masks induced differences in non-judgment processes—that is, the stimulus encoding/sensory input and response execution/motor output processes that precede or follow, respectively, the evidence accumulation process (Lerche & Voss, 2018). Masks reversed the sign of the differences in non-judgment processes across all three emotional expressions, such that judgments were faster for the inaccurate response for fearful and happy faces. Because of the inherent ambiguity to surprised expressions, there is no accurate response per se; however, masks still induced a reversal of the sign of the differences in non-judgments processes, such that judgments were faster for the less frequent "positive" judgments for faces with masks relative to those without coverings. Sunglasses had no such effect on the sign of the differences in non-judgment processes relative to faces without coverings. Because of the reduced amount of available visual information in masked faces, we expect that the effect is more likely to be driven by accelerated, yet deficient, stimulus encoding/sensory input prior to the initiation of the evidence accumulation process, rather than differences in the speed of motor execution that follow evidence accumulation.

Applications and future directions

Masks induced deficits in the evidence accumulation and the stimulus encoding process, highlighting a need for compensatory strategies to maintain effective nonverbal communication. Judgments of faces ordinarily occur on remarkably fast timescales (Bar et al., 2006; Willis & Todorov, 2006), and even a relatively small delay between the actual and expected time for a facial expression to convey meaning to another might strain communication. Indeed, the slowed evidence accumulation resulting from masks may contribute to communication strains, such as reduced feelings of interpersonal closeness caused by mask-wearing (i.e., trustworthiness, likability; Grundmann et al., 2021). That said, erroneous judgments (e.g., interpreting your neighbor's smile as a scowl) likely cause significantly more strain than a delayed response. Thus, there is a need for methods to reduce such erroneous judgments. One method for overcoming undesirable mask-related interference is instructing individuals to modulate their own response cautiousness (i.e., increasing the distance of the judgment threshold, *a*) to help counteract fast errors. Such a strategy may be particularly important to implement in contexts where individuals experience challenges with nonverbal communication (e.g., among individuals with social anxiety disorder, autism spectrum disorders; Kleberg et al., 2017; Pazhoohi et al., 2021; Saint & Moscovitch, 2021).

A more intractable aspect of mask interference is the effect on non-judgment processes (i.e., stimulus encoding/sensory input and response execution/motor output time). We expect the effect is primarily due to reduced information and faster stimulus encoding for faces with masks. Because it is unclear how one might facilitate or enhance stimulus encoding without altering the stimulus itself (e.g., removing the mask), it may be challenging to mitigate differences in non-judgment processes from the perceiver's viewpoint. That said, depending on the duration of widespread mask use, people (expressers) may begin to show adaptive changes like expressing more with the upper half of the face or using more body language to convey emotion (e.g., posture; Dael et al., 2012; Mheidly et al., 2020). Because either strategy should enrich the information conveyed by the expresser, it could offset the mask-related differences in non-judgment processes among perceivers. Future work could inform the degree to which expresser-generated strategies impact the judgment process in perceivers (e.g., by experimentally manipulating expression intensity). Alternatively, it could be that the differences in non-judgment processes are attributable to prior experience. In fact, the lack of meaningful effect sizes for the sunglasses, which also obscure facial information and could plausibly affect stimulus encoding/sensory input, suggests that extensive experience may mitigate differences in non-judgment processes. Longitudinal studies would be best suited to answer such a question. Thus, there remain questions for future research that could inform which kind of strategies are the most appropriate solutions (e.g., expresser-generated strategies vs. manipulating experience-dependent processes).

Conclusion

The present findings make a novel contribution to the literature on facial expressions of emotion by leveraging valence judgments, rather than expression recognition, to further characterize the impacts of ecologically valid face coverings. We found that masks led to more erroneous judgments of faces conveying clear valence and slowed RTs compared to faces without coverings. Drift diffusion modeling suggests that a slowed rate of evidence accumulation coupled with differences in non-judgment processing (e.g., stimulus encoding/sensory input) underlie mask-related interference. This work provides future directions for research and applied settings aimed at mitigating these effects (e.g., instructing perceivers to exercise caution and prioritize accuracy in interactions, instructing expressers to emphasize upper facial signals), which may be particularly beneficial for some individuals (e.g., social anxiety and autism spectrum disorders).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Institutes of Health under Grant NIMH111640 (PI: Neta) and Grant NIDA022975 (T32 Trainee: Harp), and the National Science Foundation under Grants CAREER Award #1752848 (PI: Neta) and RAPID Award, #2031101 (PI: Neta).

Data availability statement

Data are available on the Open Science Framework (https://osf.io/2wn7r/)

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*(2), 269–278. https://doi.org/10.1037/1528-3542.6.2.269

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest: a Journal of the American Psychological Society*, *20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Basyouni, R., Harp, N. R., Haas, I. J., & Neta, M. (2022). Political identity biases Americans' judgments of outgroup emotion. *Journal of Experimental Social Psychology*, *103*, 104392. https://doi.org/10.1016/j.jesp.2022.104392

Brock, R. L., Harp, N. R., & Neta, M. (2022). Interpersonal emotion regulation mitigates the link between trait neuroticism and a more negative valence bias. *Personality and Individual Differences*, *196*, 111726. https://doi.org/10.1016/j.paid.2022.111726

Caniato, M., Marzi, A., & Gasparella, A. (2021). How much COVID-19 face protections influence speech intelligibility in classrooms? *Applied Acoustics. Acoustique Applique. Angewandte Akustik*, *178*, 108051. https://doi.org/10.1016/j.apacoust.2021.108051

Carbon, C.-C. (2020). Wearing face masks strongly confuses counterparts in reading emotions. *Frontiers in Psychology*, *11*, 566886. https://doi.org/10.3389/fpsyg.2020.566886

Carbon, C.-C., & Serrano, M. (2021). the impact of face masks on the emotional reading abilities of children—a lesson from a joint school–university project. *i-Perception*, *12*(4), 20416695211038265. https://doi.org/10.1177/20416695211038265

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, *135*(2), 183–204. https://doi.org/10.1037/a0013965

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4–19. https://doi.org/10.1016/j.jesp.2015.07.006

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion (Washington, D.C.)*, *12*(5), 1085–1101. https://doi.org/10.1037/a0025737

Du, S., & Martinez, A. M. (2015). Compound facial expressions of emotion: From basic research to clinical applications. *Dialogues in Clinical Neuroscience*, *17*(4), 443–455. https://doi.org/10.31887/DCNS.2015.17.4/sdu

Ekman, P., & Friesen, W. V. (1978). Facial action coding system (FACS) [Database record]. APA PsycTests.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review*, *105*(3), 482–498. https://doi.org/10.1037/0033-295x.105.3.482

Fischer, A. H., Gillebaart, M., Rotteveel, M., Becker, D., & Vliek, M. (2012). Veiled emotions: the effect of covered faces on emotion perception and attitudes. *Social Psychological and Personality Science*, *3*(3), 266–273. https://doi.org/10.1177/1948550611418534

Gelfand, M., Li, R., Stamkou, E., Pieper, D., Denison, E., Fernandez, J., Choi, V., Chatman, J., Jackson, J., & Dimant, E. (2022). Persuading republicans and democrats to comply with mask wearing: An intervention tournament. *Journal of Experimental Social Psychology*, *101*, 104299. https://doi.org/10.1016/j.jesp.2022.104299

Gibson, B. M., Wasserman, E. A., Gosselin, F., & Schyns, P. G. (2005). Applying bubbles to localize features that control pigeons' visual discrimination behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*, *31*(3), 376–382. https://doi.org/10.1037/0097-7403.31.3.376

Gori, M., Schiatti, L., & Amadeo, M. B. (2021). Masking emotions: face masks impair how we read emotions. *Frontiers in Psychology*, *12*, 669432. https://doi.org/10.3389/fpsyg.2021.669432

Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. *PloS One*, *16*(4), e0249792. https://doi.org/10.1371/journal.pone.0249792

Harmon-Jones, E., & Allen, J. J. B. (1998). Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology*, *74*(5), 1310–1316. https://doi.org/10.1037//0022-3514.74.5.1310

Harp, N. R., Brown, C. C., & Neta, M. (2021). Spring break or heart break? extending valence bias to emotional words. *Social Psychological and Personality Science*, *12*(7), 1392–1401. https://doi.org/10.1177/1948550620972296

Harp, N. R., Freeman, J. B., & Neta, M. (2022). Mindfulness-based stress reduction triggers a long-term shift toward more positive appraisals of emotional ambiguity. *Journal of Experimental Psychology. General*, *151*(9), 2160–2172. https://doi.org/10.1037/xge0001173

Howard, J., Huang, A., Li, Z., Tufekci, Z., Zdimal, V., van der Westhuizen, H.-M., von Delft, A., Price, A., Fridman, L., Tang, L.-H., Tang, V., Watson, G. L., Bax, C. E., Shaikh, R., Questier, F., Hernandez, D., Chu, L. F., Ramirez, C. M., & Rimoin, A. W. (2021). An evidence review of face masks against COVID-19. *Proceedings of the National Academy of Sciences*, *118*(4), e2014564118. https://doi.org/10.1073/pnas.2014564118

Johnson, D. J., Hopwood, C. J., Cesario, J., & Pleskac, T. J. (2017). Advancing research on cognitive processes in social and personality psychology: a hierarchical drift diffusion model primer. *Social Psychological and Personality Science*, *8*(4), 413–423. https://doi.org/10.1177/1948550617703174

Kim, G., Seong, S. H., Hong, S.-S., & Choi, E. (2022). Impact of face masks and sunglasses on emotion recognition in South Koreans. *PloS One*, *17*(2), e0263466. https://doi.org/10.1371/journal.pone.0263466

Kleberg, J. L., Högström, J., Nord, M., Bölte, S., Serlachius, E., & Falck-Ytter, T. (2017). Autistic traits and symptoms of social anxiety are differentially related to attention to others' eyes in social anxiety disorder. *Journal of Autism and Developmental Disorders*, *47*(12), 3814–3821. https://doi.org/10.1007/s10803-016-2978-z

Krieglmeyer, R., Deutsch, R., De Houwer, J., & De Raedt, R. (2010). Being moved: Valence activates approach-avoidance behavior independently of evaluation and approach-avoidance intentions. *Psychological Science*, *21*(4), 607–613. https://doi.org/10.1177/0956797610365131

Langbehn, A. T., Yermol, D. A., Zhao, F., Thorstenson, C. A., & Niedenthal, P. M. (2022). Wearing N95, surgical, and cloth face masks compromises the perception of emotion. *Affective Science*, *3*(1), 105–117. https://doi.org/10.1007/s42761-021-00097-z

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition & Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lerche, V., & Voss, A. (2018). Speed–accuracy manipulations and diffusion modeling: Lack of discriminant validity of the manipulation or of the parameter estimates? *Behavior Research Methods*, *50*(6), 2568–2585. https://doi.org/10.3758/s13428-018-1034-7

Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods*, *49*(2), 513–537. https://doi.org/10.3758/s13428-016-0740-2

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces—KDEF, CD ROM*. Karolinska Institutet. https://www.kdef.se/home/aboutKDEF.html

Mheidly, N., Fares, M. Y., Zalzale, H., & Fares, J. (2020). Effect of face masks on interpersonal communication during the COVID-19 pandemic. *Frontiers in Public Health*, *8*, 582191. https://doi.org/10.3389/fpubh.2020.582191

Neta, M., & Brock, R. L. (2021). Social connectedness and negative affect uniquely explain individual differences in response to emotional ambiguity. *Scientific Reports*, *11*(1), 1. https://doi.org/10.1038/s41598-020-80471-2

Neta, M., & Dodd, M. D. (2018). Through the eyes of the beholder: Simulated eye-movement experience ("SEE") modulates valence bias in response to emotional ambiguity. *Emotion*, *18*(8), 1122–1127. https://doi.org/10.1037/emo0000421

Neta, M., & Tong, T. T. (2016). Don't like what you see? Give it time: Longer reaction times associated with increased positive affect. *Emotion*, *16*(5), 730–739. https://doi.org/10.1037/emo0000181

Neta, M., Davis, F. C., & Whalen, P. J. (2011). Valence resolution of ambiguous facial expressions using an emotional oddball task. *Emotion*, *11*(6), 1425–1433. https://doi.org/10.1037/a0022993

Neta, M., Harp, N. R., Henley, D. J., Beckford, S. E., & Koehler, K. (2019). One step at a time: Physical activity is linked to positive interpretations of ambiguity. *PLoS One*, *14*(11), e0225106. https://doi.org/10.1371/journal.pone.0225106

Neta, M., Kelley, W. M., & Whalen, P. J. (2013). Neural responses to ambiguity involve domain-general and domain-specific emotion processing systems. *Journal of Cognitive Neuroscience*, *25*(4), 547–557. https://doi.org/10.1162/jocn_a_00363

Neta, M., Norris, C. J., & Whalen, P. J. (2009). Corrugator muscle responses are associated with individual differences in positivity-negativity bias. *Emotion*, *9*(5), 640–648. https://doi.org/10.1037/a0016819

Neta, M., Tong, T. T., Rosen, M. L., Enersen, A., Kim, M. J., & Dodd, M. D. (2017). All in the first glance: First fixation predicts individual differences in valence bias. *Cognition & Emotion*, *31*(4), 772–780. https://doi.org/10.1080/02699931.2016.1152231

Niedenthal, P. M., Hampton, R. S., & Marji, M. (2023). Ancestral Diversity: A Socioecological Account of Emotion Culture. *Current Directions in Psychological Science*, *32*(2), 167–175. https://doi.org/10.1177/09637214221151154

Niedenthal, P. M., Rychlowska, M., Wood, A., & Zhao, F. (2018). Heterogeneity of long-history migration predicts smiling, laughter and positive emotion across the globe and within the United States. *PLoS One*, *13*(8), e0197651. https://doi.org/10.1371/journal.pone.0197651

Noyes, E., Davis, J. P., Petrov, N., Gray, K. L. H., & Ritchie, K. L. (2021). The effect of face masks and sunglasses on identity and expression recognition with super-recognizers and typical observers. *Royal Society Open Science*, *8*(3), 201169. https://doi.org/10.1098/rsos.201169

Pazhoohi, F., Forby, L., & Kingstone, A. (2021). Facial masks affect emotion recognition in the general population and individuals with autistic traits. *PLoS One*, *16*(9), e0257740. https://doi.org/10.1371/journal.pone.0257740

Pleskac, T. J., Cesario, J., & Johnson, D. J. (2018). How race affects evidence accumulation during the decision to shoot. *Psychonomic Bulletin & Review*, *25*(4), 1301–1330. https://doi.org/10.3758/s13423-017-1369-6

R Core Team. (2022). *R (4.2.1). R foundation for statistical computing*. https://www.R-project.org/

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology. Human Perception and Performance*, *26*(1), 127–140. https://doi.org/10.1037//0096-1523.26.1.127

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, *16*(2), 323–341. https://doi.org/10.1037/0882-7974.16.2.323

Roberson, D., Kikutani, M., Döge, P., Whitaker, L., & Majid, A. (2012). Shades of emotion: What the addition of sunglasses or masks to faces reveals about the development of facial expression processing. *Cognition*, *125*(2), 195–206. https://doi.org/10.1016/j.cognition.2012.06.018

Ruba, A. L., & Pollak, S. D. (2020). Children's emotion inferences from masked faces: Implications for social interactions during COVID-19. *PLoS One*, *15*(12), e0243708. https://doi.org/10.1371/journal.pone.0243708

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295x.110.1.145

Rychlowska, M., Korb, S., Brauer, M., Droit-Volet, S., Augustinova, M., Zinner, L., & Niedenthal, P. M. (2014). Pacifiers disrupt adults' responses to infants' emotions. *Basic and Applied Social Psychology*, *36*(4), 299–308. https://doi.org/10.1080/01973533.2014.915217

Rychlowska, M., Miyamoto, Y., Matsumoto, D., Hess, U., Gilboa-Schechtman, E., Kamble, S., Muluk, H., Masuda, T., & Niedenthal, P. M. (2015). Heterogeneity of long-history migration explains cultural differences in reports of emotional expressivity and the functions of smiles. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(19), E2429–E2436. https://doi.org/10.1073/pnas.1413661112

Saint, S. A., & Moscovitch, D. A. (2021). Effects of mask-wearing on social anxiety: An exploratory review. *Anxiety, Stress & Coping: An International Journal, No Pagination Specified-No Pagination Specified*, *34*(5), 487–502. https://doi.org/10.1080/10615806.2021.1929936

Samuelsson, H., Jarnvik, K., Henningsson, H., Andersson, J., & Carlbring, P. (2012). The Umeå University Database of Facial Expressions: A validation study. *Journal of Medical Internet Research*, *14*(5), e136. https://doi.org/10.2196/jmir.2196

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*(5), 402–409. https://doi.org/10.1111/1467-9280.00472

Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*(1), 101–117. https://doi.org/10.1037/0278-7393.32.1.101

Theisen, M., Lerche, V., von Krause, M., & Voss, A. (2021). Age differences in diffusion model parameters: A meta-analysis. *Psychological Research*, *85*(5), 2012–2021. https://doi.org/10.1007/s00426-020-01371-8

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Trafimow, D., & Marks, M. (2015). Editorial. *Basic and Applied Social Psychology*, *37*(1), 1–2. https://doi.org/10.1080/01973533.2015.1012991

Trafimow, D., & Osman, M. (2022). Barriers to converting applied social psychology to bettering the human condition. *Basic and Applied Social Psychology*, *44*(1), 1–11. https://doi.org/10.1080/01973533.2022.2051327

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767–775. https://doi.org/10.3758/bf03192967

Voss, A., Rothermund, K., & Brandtstädter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, *44*(4), 1048–1056. https://doi.org/10.1016/j.jesp.2007.10.009

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206–1220. https://doi.org/10.3758/bf03196893

Voss, A., Voss, J., & Klauer, K. C. (2010). Separating response-execution bias from decision bias: Arguments for an additional parameter in Ratcliff's diffusion model. *The British Journal of Mathematical and Statistical Psychology*, *63*(Pt 3), 539–555. https://doi.org/10.1348/000711009X477581

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology*, *6*, 336. https://doi.org/10.3389/fpsyg.2015.00336

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A Diffusion Model Account of Criterion Shifts in the Lexical Decision Task. *Journal of Memory and Language*, *58*(1), 140–159. https://doi.org/10.1016/j.jml.2007.04.006

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Willis, J., & Todorov, A. (2006). First impressions: making up your mind after a 100-Ms exposure to a face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Ziccardi, S., Crescenzo, F., & Calabrese, M. (2021). What is hidden behind the mask?" facial emotion recognition at the time of COVID-19 pandemic in cognitively normal multiple sclerosis patients. *Diagnostics 12*(1), 47. https://doi.org/10.3390/diagnostics12010047