Drift-Diffusion Modeling Reveals that Masked Faces are Preconceived as Unfriendly

Martijn J. Mulder (m.j.mulder@uu.nl)
Utrecht University

Franziska Prummer
Lancaster University

David Terburg
Utrecht University

J. Leon Kenemans
Utrecht University

Article

Keywords:

Posted Date: June 17th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1744487/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 9th, 2023. See the published version at https://doi.org/10.1038/s41598-023-44162-y.

Abstract

During the COVID-19 pandemic, the use of face masks has become a daily routine. Studies have shown that face masks increase the ambiguity of facial expressions which not only affects (the development of) emotion recognition, but also interferes with social interaction and judgement. To disambiguate facial expressions, we rely on *perceptual* (stimulus-driven) as well as *judgmental* (preconception) processes. However, it is unknown which of these two mechanisms accounts for the misinterpretation of masked expressions. To investigate this, we asked participants (N = 136) to decide whether ambiguous (morphed) facial expressions, with or without a mask, were perceived as friendly or unfriendly. To test for the independent effects of *perceptual* and *judgmental* biases we fitted a drift-diffusion model (DDM) to the behavioral data of each participant. Results show that face masks induce a clear loss of information leading to a slight *perceptual* bias towards friendly choices, but also a clear *judgmental* bias towards unfriendly choices for masked faces. These results suggest that, although face masks can increase the perceptual friendliness of faces, people have the prior preconception to interpret masked faces as unfriendly.

Introduction

During the COVID-19 Pandemic, wearing a face mask has become part of our daily life as it restricts the spread of the SARS-CoV-2 virus[1,2]. Since facial expressions play an important role in our social communication[3–6], wearing such a mask might also affect social conduct. For example, when we accidently step on someone's toes in the supermarket, apologizing with a friendly smile might not be sufficient to save the situation. Indeed, recent studies show that facial masks affect face perception, recognition and identification[7–14,14–16], interfere with social interaction and social judgments[7,13,14] and might even hamper the development of emotion recognition in children[12,17].

Importantly, face masks not only reduce the amount of sensory information, but potentially also influence the classification of facial expression in a more systematically biased way by adding perceptual cues (e.g., masks could make people look more angry or sad), and/or evoking preconceptions about obscured emotional expressions. In other words, the detrimental effects face masks have on our social interaction[7,13,14] can be due to perceptual aspects of emotion recognition, but also due to top-down preconceived judgement effects. Given the general social importance of facial emotion recognition it is vital to disentangle such perceptual from judgmental biases. Especially since face masks have become part of our daily routine, awareness of how this can impact our social communication might help to counter the potential negative consequences. Whether perceptual and/or judgmental biases account for the misinterpretation of masked facial expressions is however currently unknown.

Evidence showing misinterpretation of emotional expressions due to face masks lies in agreement with previous research into the effects of occlusion of the lower part of the face in several emotional expressions (e.g. [18–21]). This especially holds for the identification of happy expressions, for which people rely more on the mouth-region; in contrast, for identifying angry expressions, the eye-region seems

to be the most prominent diagnostic cue[21–26]. However, most of the studies on the effects of facial masks used facial stimuli with full prototypical emotional expressions (e.g., happy, angry, surprise, fear, sadness, disgust), ignoring the fact that emotional expressions during our daily life are often less intense and not profoundly demarcated. As such, facial expressions are often ambiguous, making their interpretation more susceptible to a perceptual (stimulus driven) or judgmental (preconception) bias[21,27–34]. Given this ambiguity of emotional expressions in daily life, the question arises how facial masks affect the interpretation of facial expressions. For example, the occlusion of a moderately friendly smile might result in an interpretation of the expression based on the eyes only, which might especially be problematic when the smile is not completely sincere and only used as a social gesture or even used to mask negative feelings[35,36]. In this sense occlusion of the mouth results in a loss of sensory information which may result in a *perceptual (stimulus-driven) bias* away from smile-driven friendliness[37,38].

On the other hand, as already introduced above, masks may also serve as an (unintended) perceptual cue. For instance, although wearing a facial mask during the Covid-19 pandemic is mostly accepted[9], facial masks can still elicit a negative association due to occlusion of important parts of the face[9,39–43], which might in turn result in a tendency to interpret an ambiguous emotional expression as a negative, unfriendly appearance. Such contextual (goal-directed) effects have proven to affect the interpretation of emotional expression as well, resulting in a *judgmental (top-down) bias*[27,28,38].

In sum, both the perceptual (stimulus-driven) and the judgmental (top-down) perspectives, predict that masks will elicit a stronger tendency (bias) to classify facial expressions more often as negative (e.g., unfriendly). To investigate whether facial masks elicit such perceptual and/or judgmental biases in the interpretation of ambiguous emotional expressions, we conducted an experiment in which participants were asked whether ambiguous happy or angry expressions, with and without facial masks, are perceived as friendly or unfriendly.

Perceptual and judgmental biases are hard to separate, as both biases result in faster and more choices for a favored alternative. To distinguish between a possible perceptual or a judgmental bias, we use the drift diffusion model (DDM)[44–46]. The drift-diffusion model allows to decompose the underlying choice process and quantify a possible perceptual or judgmental bias by utilizing both accuracy and reaction time data[46–48]. The model assumes that, during a perceptual choice, noisy sensory evidence accumulates until a decision threshold is hit (Fig. 1B; for reviews see refs. [45,48–52]). For instance, when a facial mask affects the uptake of *friendly* information (e.g., smile) at the stimulus level, the accumulation process will change in favor for the unfriendly alternative, resulting in a perceptual bias, with faster and more ‘unfriendly’ choices (see Fig. 1C). This process is thus sensitive to *stimulus-driven* biases, but at the same time the starting-point of evidence accumulation can differ based on *top-down* priors. For instance, a negative top-down effect will affect the symmetry between the two decision thresholds, resulting in faster and more ‘unfriendly’ responses as well, as the threshold for the amount of ‘unfriendly’ information will be lower than the threshold for ‘friendly’ information (see Fig. 1D). Even when these biases result in similar behavioral changes, drift-diffusion modelling will allow us to disentangle them and answer the

question whether the face-mask driven impairments in the interpretation of ambiguous facial expressions are due to perceptual and/or judgmental biases.

To test for a possible perceptual or judgmental bias, we fitted the DDM to participants' performance on judging emotionally ambiguous faces, with and without mask, on friendliness. Overall, we expect lower drift-rates for masked facial expressions, reflecting less sensory information to make the correct decision, resulting in slower and more error prone choices. In addition, our methodology allows to disentangle a perceptual from a judgmental bias in the assessment of the friendliness of masked and unmasked ambiguous facial expressions. We expect that if such a judgmental bias is indeed present when assessing masked faces, the distance between the start and end point of the accumulation process will be smaller for the unfriendly compared to friendly alternative, due to the negative connotations typically associated with face masks [40,42,43]. Furthermore, we will test whether facial masks induce a perceptual (stimulus-driven) bias due to the occlusion of the mouth region and possibly due to the diagnostic cues in the visual features of the mask itself.

Results

Below we will first report the effect of stimulus ambiguity and mask on choice and response time data. Next, we will disentangle perceptual and judgmental biases using DDM analyses that explain the descriptive results in terms of parameter changes.

Descriptive results

To quantify the effect of facial mask on the interpretation of ambiguous facial expressions, a logistic function was fit on the choice data (Eq. 1; Fig. 2A). For both masked and unmasked ambiguous facial expressions, the proportion unfriendly choices increased as a function of stimulus ambiguity (from happy to angry: see Fig. 2A). For masked faces, there was a small, but significant negative choice bias (β0) reflecting a tendency to choose for friendly more often ($b0_{mdn}$= -0.22, one-sample Wilcoxon signed-rank test for $b0_{mdn}$ = 0, V = 2793, $p < 0.01$). No significant bias was found for unmasked faces. Sensitivity (b1) to the stimulus was significantly lower for masked ($b1_{mdn}$ = 6.47) vs unmasked ($b1_{mdn}$ = 10.91) facial expressions (Wilcoxon signed-rank test, W = 9151, $p < 0.01$).

We tested for significant effects in response times using a 2 (masked vs. unmasked) × 6 (stimulus ambiguity) repeated measures analysis of variance (ANOVA). For response times, there was a significant main effect of emotional ambiguity of the facial expression, with increasing response times for lower ambiguity levels, symmetrical around zero intensity, $F(1,135) = 366.4$, $p < 0.01$. The main effect of mask was significant as well, with slower response times for masked stimuli, $F(5,675) = 290.5$, $p < 0.01$. In addition, there was a significant interaction effect between the stimulus ambiguity of the facial expression and mask, indicating that the effect of mask was not equally distributed across ambiguity levels, $F(5,675) = 44.1$, $p < 0.01$. More specifically, the difference between response times for masked and no-masked facial expression became smaller for the low (-10, 10) ambiguity levels (see Fig. 2B).

Furthermore, post-hoc t-test show significant slower response times for masked happy than for masked angry facial expressions with a high (-60 vs 60) or moderate (-40 vs 40) stimulus ambiguity (both *ts*(135) > 5.4, $p < 0.01$). No such difference was found for the facial stimuli without a mask. Instead, participants were slower for angry vs happy facial expressions without a mask, with low emotional ambiguity (10, vs -10), $t(135) = 3.82$, $p < 0.01$.

Overall, these results of the analyses of choice and response times show that there are small, asymmetrical effects of a facial mask on interpretation of ambiguous emotional expressions. To further quantify these effects, we fitted the DDM to the data allowing us to decompose the effects in the underlying choice parameters.

DDM analyses.

The RT results in Fig. 2 partly suggest a bias towards unfriendly choices for masked stimuli, showing faster choices for easy (60) and moderate (40) angry masked faces. In contrast, the psychometric data reflect a general loss of sensitivity combined with an unexpected bias to friendly choices in the mask condition. These contradictory findings suggest that facial masks might affect the interpretation of ambiguous emotional faces via different underlying mechanisms. To identify whether bias effects are driven by a judgmental (top-down) or perceptual (stimulus-driven) process, the diffusion model was fitted to both the RT and choice data simultaneously, allowing to disentangle these different types of bias.

For the diffusion-model fits (see Fig. 4 in the methods section for a graphical representation of the goodness of fit), we found that both starting points for masked and unmasked stimuli were different from 0.5, with a median starting point of $z_{Mdn(SD)\_masked} = 0.53(0.07)$ and $z_{Mdn(SD)\_unmasked} = 0.47(0.05)$ respectively (Wilcoxon signed-rank test, $V > 1490$, $p < 0.01$), indicating a significant judgmental bias towards the unfriendly alternative for masked and towards the friendly alternative for unmasked facial expressions (see Fig. 3A).

Drift-rate varied with stimulus ambiguity, with higher (more negative for happy and more positive for angry) drift-rates for both masked and unmasked emotional expressions, $F(1, 135) = 43.4$, $p < 0.01$ (see Fig. 3B). In general, drift-rates were lower (less positive for angry and less negative for happy) for masked facial expressions, indicated by the significant interaction between stimulus ambiguity and mask, $F(1,135) = 92.9$, $p < 0.01$. To test for a perceptual bias, we first tested for a linear relationship between (signed) stimulus ambiguity and drift-rate for masked and unmasked facial expressions separately. Regression analyses showed that stimulus ambiguity predicted drift-rate for masked ($R^2 = .86$, $F(1, 815) = 4891$, $p < .001$) and unmasked ($R^2 = 0.90$, $F(1, 815) = 7193$, $p < 0.01$) facial stimuli. Next, perceptual bias was defined as the vertical shift (e.g., intercept) of the regression lines, representing a possible systematic bias towards either the friendly or unfriendly alternative (see Fig. 3B). The intercept was − 0.068 ($t = -2.49$, $p < 0.01$) for masked and 0.25 ($t = 8.48$, $p < 0.01$) for unmasked facial expressions, suggesting a small perceptual bias towards friendly choices for masked stimuli and a small perceptual bias towards unfriendly for unmasked facial expressions.

In addition to starting point (z) and drift-rate (v), non-decision time (*t0*) was free to vary across masked and unmasked conditions as well, to account for possible differences in sensory encoding of the stimulus, prior to the decision process. Non-decision time (*t0*) was higher for masked ($t0_{Mdn(SD)}$ = 0.51(0.06)) compared to unmasked facial expressions ($t0_{Mdn(SD)}$ = 0.47(0.05)), presumably reflecting a longer time for perceptual encoding, prior to the accumulation process (Wilcoxon signed-rank test, $W = 228$, $p < 0.01$).

In sum, our descriptive analyses suggest that masked compared to unmasked faces are judged as more friendly, but that judging masked friendly faces takes more time. Our DDM analyses show that masking a face results in a loss of sensory information and an unfriendly judgmental bias. In contrast, we found a slight friendly perceptual bias as well. This suggests that, although diagnostic cues in masked faces bias our participants towards friendliness via a stimulus-driven process, our participants also have the preconception that masked faces are unfriendly.

Discussion

To investigate whether face masks induce a perceptual or judgmental bias in the interpretation of ambiguous facial expressions, participants performed a task in which they had to decide whether ambiguous facial expressions, with or without a mask, were perceived as friendly or unfriendly. We fitted a drift-diffusion model (DDM) to their performance data to test for the independent effects of perceptual and judgmental biases in these decisions. As expected, the analyses of descriptive data showed a lower sensitivity for the masked emotional expressions, generally resulting in slower and less correct responses for the masked compared to the unmasked facial expressions. This effect is in line with studies showing that covering the mouth decreases the amount of available information to correctly recognize and identify an emotional expression[7–14,14–16,18–21].

In addition, quantification of choice data using the psychometric function shows a small but significant bias toward *friendly* faces for masked but not for unmasked facial expressions. This is unexpected, since the mouth is often considered to be more important in the recognition of happy facial expressions compared to the recognition of angry facial expressions[21–26]. As such, based on perceptual processes alone, we expected that wearing a facial mask would especially hamper the identification of happy facial expressions, resulting in a bias *away* from the friendly choices for masked stimuli. Instead, the small bias towards friendly choices suggests that covering the mouth with a facial mask has a larger effect on the misinterpretation of angry than happy facial expressions, which seems to be particularly the case for expressions with a high emotional ambiguity (i.e., -10 and 10; see Fig. 2A). In contrast, analyses of average response times show slower response times for happy than for angry masked facial expressions with high (-60 vs 60) or moderate (-40 vs 40) emotional ambiguity. These contradictory findings in the descriptive data underscore the importance to fit a computational model to the data that takes into account both choice and response time data, allowing us to disentangle the underlying biasing mechanisms.

To measure possible systematic perceptual (stimulus driven) or judgmental (top-down) biases in the interpretation of masked and unmasked facial expressions, we fitted the drift-diffusion model to each participant's choice and response time data. Results show that facial masks affect both the perceptual and judgmental processes, in opposite direction, with a judgmental bias towards unfriendly and a perceptual bias towards friendly choices.

As expected, we found a small but significant shift in the starting point towards the unfriendly alternative. This bias in starting point suggests that participants start the decision process with asymmetrical bounds (i.e., smaller for unfriendly vs friendly, in masked choices), resulting in faster and more choices for the unfriendly alternative. Such a lower bound might indicate a top-down preconception in which the alternatives already have a different representation for masked versus unmasked stimuli, prior to the initial choice. This might be due to the somewhat threatening connotation of the mask, providing a context which might bias the interpretation of ambiguous facial expressions[27,28].

Additionally, we found a judgmental bias towards the friendly bound for the *unmasked* facial stimuli. This could imply that our participants have the preconception that emotionally ambiguous faces are friendly, but this positive bias could also reflect a reversed effect on faces when the mask is *not* visible, resulting in a more positive connotation than usual. Next to the judgmental (top-down) bias for masked faces, we found a perceptual (stimulus driven) bias towards *friendly* sensory information for masked stimuli. One explanation for this unexpected perceptual bias may be related to diagnostic features of the mask itself. Studies investigating the impact of the emotional intensity of the facial expression show that happy expressions are more easily detected, even at low intensities[54] and resolutions[55]. Given the difference in the detectability of happy vs angry at a low emotional intensity, it might be the case that the low-level visual features of the mask are closer to a happy than an angry mouth expression. This might particularly be the case for early perceptual processes that are primarily affected by low-level visual features[38,56,57] in which the mask-features might result in a surrogate smile for the face, biasing the effect away from unfriendliness. As such, the drift-rates might be biased towards sensory evidence in favor for the *friendly* alternative for masked faces, due to the asymmetry in happy/angry sensitivity, where the effect of *happy-information* in the mask itself is stronger than the angry diagnostic features in the eyes of masked faces. Furthermore, it has been shown that low spatial frequencies of facial expressions are faster and earlier processed in the information stream, compared to high spatial frequencies[38,56,58,59]. In light of this explanation, our choices might be biased first by the outstanding low-level spatial features (stimulus-driven), while the semantic (top-down) categorization is processed later in time[56,57].

Alternatively, the small shift towards unfriendly choices for unmasked faces might suggest that ambiguity is not symmetrically distributed across the emotional dimension (happy to angry) for our facial stimuli. This asymmetry seems in turn to disappear after adding a mask to the faces, suggesting that this asymmetry is driven by the mouth region. Future research that includes a condition with a mouth that is covered, but not by a mask, can resolve this issue by showing whether the perceptual bias towards

friendliness is due to the additive effect of the mask, or due to the reduction of asymmetry in ambiguity by covering the mouth.

In sum, we investigated whether face masks induce a loss of information and perceptual or judgmental biases, participants were asked to decide whether masked or unmasked ambiguous facial expressions were perceived as friendly or unfriendly. Results show that wearing a face mask causes a loss in sensory information and a judgmental (preconception) bias towards *unfriendly* but a perceptual (stimulus-driven) bias towards *friendly* choices for masked faces. These results suggest that people have a prior top-down tendency to interpret facial masks as unfriendly, regardless of the friendly (stimulus-driven) effects of the facial mask itself.

Methods

Participants

Participants (n = 145, mean(std) age = 22.3(4.4), 109 female) were invited via online media or Utrecht University's Sona Systems (https://www.sona-systems.com/) to participate in an online experiment in exchange for course credit. Nine participants were excluded based on insufficient performance on the task (see descriptive analyses below). Informed consent was obtained from all participants. The experiment was approved by and was in accordance with the guidelines and ethical standards of the the Ethics Committee of Utrecht University (EC-FETC18-129).

Materials and Stimuli

Face stimuli were adapted from the Averaged Karolinska Directed Emotional Faces (AKDEF[53]). First, to control for possible sex differences in facial expressions ([60,61]), we created an angry and a happy *non-binary* face by morphing the average (resp. angry and happy) male and female faces to each other using WinMorph (version 3.01). Next, emotionally ambiguous faces were created by morphing the happy face towards the angry face in 41 incremental steps of 2.5% each. From this range of morphed non-binary facial stimuli with different angry/happy ratios, eight ambiguous expressions were chosen. For each face, a masked version was created, by adding surgical mask to each face using Adobe Photoshop (version 22.2). A normal surgical mask was chosen, as these were seen commonly in public at the time of data collection. The color of the masks was adjusted to resemble the black-and-white coloring of the images of the faces. The edges of the mask were softened to incorporate it more naturally into the image.

Six of the facial expressions with ratios (angry/happy%) of 80/20%, 70/30%, 55/45%, 45/55%, 30/70% and 20/80%, with and without a mask, were used as main stimuli (see Fig. 1B). Two facial expressions (60/40% and 40/60%), with and without a mask served as filler trials to add more variance to the stimuli, reducing predictability of the six main facial stimuli. Stimulus ambiguity was expressed as the difference between the percentage happy and angry facial expressions (assuming 50/50 to have 0% evidence for either a happy or angry expression and thus full ambiguity) resulting in 6 (signed) ambiguity levels of -60, -40, -10, 10, 40 and 60% ordered from more happy to more angry facial expression.

Procedure

A two-alternative forced (2AFC) choice task was set up and hosted on Gorilla Experiment Builder (www.gorilla.sc[62]). After consent was given, general demographic information was collected after which the participant was assigned to one of the two (counterbalanced) versions of the 2AFC-task. In the 2AFC-task, participants were asked to respond as quickly as possible to decide whether the facial expression was perceived as friendly or unfriendly, for a total of 608 trials. These trials consist of 96 trials (48 masked) for each of the six stimulus ambiguities (-60%, -40%, -10%, 10%, 40%, 60%) and 32 (16 masked) filler-trials (ambiguities – 20% and 20%). To keep the participants engaged, the experiment was divided into 8 blocks of 76 trials each. Each block contained a random alternation of all possible conditions (mask x stimulus ambiguity).

Each trial started with a fixation cross that was presented for a randomly chosen duration between 600 and 1200ms to prevent anticipatory responses to the stimulus. Next, the stimulus was shown on the screen, during which the participant was required to respond with the 'C'- or 'M'-key to indicate their choice. Stimulus display was terminated after a button press or a time-out of 2300ms. Choice associations with these responses ('friendly' or 'unfriendly') were counterbalanced between participants. We chose to use the labels *friendly* and *unfriendly* since the created images were not fully 'angry' or 'happy'. For example, the ambiguous facial expression with 10% stimulus strength might not be perceived as angry perse, but still has a mild 'unfriendly' expression. Subsequently, the participant's response feedback was shown for 400 ms (a green check for correct and a red cross for false responses). Whenever a response was made throughout the fixation cross period, an icon with the words "too fast" appeared. If subjects did not make a response within the given response time (2300 ms), the word "miss" was shown.

Analyses

Descriptive data. For each participant, response times were log transformed, after which for each condition response times were removed that were three standard deviations away from the average response time (on average, 4.7% of the data). Next, median response times were calculated for each condition separately. A 2 (mask) x 2 (expression) x 6 (ambiguity) repeated measures ANOVA was used to test for effects of mask, expression, and ambiguity on response times.

To quantify effects of mask and stimulus ambiguity on choice performance, a logistic function (see Eq. 1) was fit onto the choice data for each participant.

$$p_{unfriendly} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 S)}} \quad (1)$$

This function included two terms, an ambiguity-dependent term that reflected sensitivity to the emotion stimulus and an ambiguity-independent term that reflected a choice bias towards either 'friendly' or 'unfriendly' choices. Non-parametric Wilcoxon t-tests were used to test for a difference in sensitivity (b1)

between masked and unmasked faces and for a possible choice bias (b0) in the masked and unmasked conditions. Based on these initial analyses, nine participants were excluded based on a significantly large deviance between empirical data and the fit of the psychometric function (all nine deviances > 310; X^2(df's < 280) < 242.2, $p = 0.05$).

DDM analyses. To test for a judgmental or perceptual bias, the drift-diffusion model (DDM [44]; see Fig. 1B) was fitted to each participant's choice and response time data simultaneously, using the fast-dm toolbox ([63–67]). To quantify a possible perceptual or judgmental bias, drift-rate (*v*) and starting-point (*z*) were free to vary for masked vs unmasked stimuli (see Fig. 1C, D). To capture possible differences in sensory encoding and/or motor responses, non-decision (*t0*) time was free to vary for the masked and the unmasked condition as well. In addition, drift-rate was free to vary across all 6 ambiguity levels (with negative values for the lower 'friendly' bound and positive values for the upper 'unfriendly' threshold). Decision threshold (*a*) was fixed across all conditions. Variability parameters (*sz*, *st0* and *sv*) were not fitted and set to zero as fitting these parameters can bias the estimations of the main parameters ([68,69]). The Kolmogorov–Smirnov (KS) criterion was used as optimization criterion in the model fit ([65,66]). To this end, the cumulative density distributions (CDFs) for the unfriendly (upper) and friendly (lower) bounds are combined by multiplying all RTs from the lower (friendly) bound by − 1 (see refs. [66,67]), resulting in a combined CDF for each of the twelve (mask x ambiguity) conditions (Fig. 4). During the fit process, parameter estimates are iteratively adjusted to search for the minimum distance between the empirical and predicted (model) CDF.

Goodness of fit of the resulting model was graphically visualised based on the RT distributions of all participants combined in one group-data set (with an average of 6213 trials for each of the 12 conditions). Figure 4 represents the model fit in which the cumulative distributions (CDFs) for friendly (negative sign) and unfriendly (positive sign) choices were combined for the empirical and predicted group CDFs[45]. As the Fig. indicates, the difference between the empirical and model CDFs is small.

To test for a judgmental bias in each condition, one sample Wilcoxon signed-rank tests were used to test whether starting-points z_{masked} and $z_{unmasked}$ where different from $z = 0.5$ (the value of an unbiased starting point, halfway the boundary separation *a*). In addition, to test for differences in starting-point *z* or non-decision time *t0* between the masked and unmasked conditions, paired-samples Wilcoxon signed-rank tests were used. Finally, to test for differences between masked and unmasked conditions in drift-rate, a 2 (mask vs unmasked) x 6 (ambiguity) repeated measures ANOVA was performed. Furthermore, to quantify a possible perceptual bias, we ran a regression analyses on the drift-rate values of each participant and each ambiguity level, testing for a linear relationship between (signed) stimulus ambiguity and drift-rate[48,70,71]. To test for a perceptual (stimulus-driven) bias, we tested whether the intercept of the regression lines differed significantly from zero, indicating a general tendency towards unfriendly (positive) or friendly (negative) sensory information.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Eikenberry, S. E. *et al.* To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling* **5**, 293–308 (2020).
2. Sande, M. van der, Teunis, P. & Sabel, R. Professional and Home-Made Face Masks Reduce Exposure to Respiratory Infections among the General Population. *PLOS ONE* **3**, e2618 (2008).
3. Crivelli, C. & Fridlund, A. J. Facial displays are tools for social influence. *Trends in Cognitive Sciences* **22**, 388–399 (2018).
4. Gosselin, P., Kirouac, G. & Doré, F. Y. Components and recognition of facial expression in the communication of emotion by actors. *J Pers Soc Psychol* **68**, 83–96 (1995).
5. Haxby, J. V., Hoffman, E. A. & Gobbini, M. I. Human neural systems for face recognition and social communication. *Biol Psychiatry* **51**, 59–67 (2002).
6. Nusseck, M., Cunningham, D. W., Wallraven, C. & Bülthoff, H. H. The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision* **8**, 1 (2008).
7. Calbi, M. *et al.* The consequences of COVID-19 on social interactions: an online study on face covering. *Sci Rep* **11**, 2601 (2021).
8. Carbon, C.-C. Wearing Face Masks Strongly Confuses Counterparts in Reading Emotions. *Front. Psychol.* **11**, 566886 (2020).
9. Carbon, C.-C. About the Acceptance of Wearing Face Masks in Times of a Pandemic. *i-Perception* **12**, 204166952110211 (2021).
10. Carragher, D. J. & Hancock, P. J. B. Surgical face masks impair human face matching performance for familiar and unfamiliar faces. *Cogn. Research* **5**, 59 (2020).
11. Freud, E., Stajduhar, A., Rosenbaum, R. S., Avidan, G. & Ganel, T. The COVID-19 pandemic masks the way people perceive faces. *Sci Rep* **10**, 22344 (2020).
12. Gori, M., Schiatti, L. & Amadeo, M. B. Masking Emotions: Face Masks Impair How We Read Emotions. *Front. Psychol.* **12**, 669432 (2021).
13. Grundmann, F., Epstude, K. & Scheibe, S. *Face masks reduce emotion-recognition accuracy and perceived closeness*. https://osf.io/xpzs3 (2020) doi:10.31234/osf.io/xpzs3.
14. Marini, M., Ansani, A., Paglieri, F., Caruana, F. & Viola, M. The impact of facemasks on emotion recognition, trust attribution and re-identification. *Sci Rep* **11**, 5577 (2021).
15. Nestor, M. S., Fischer, D. & Arnold, D. “Masking” our emotions: Botulinum toxin, facial expression, and well-being in the age of COVID-19. *J Cosmet Dermatol* **19**, 2154–2160 (2020).
16. Pazhoohi, F., Forby, L. & Kingstone, A. Facial masks affect emotion recognition in the general population and individuals with autistic traits. *PLoS ONE* **16**, e0257740 (2021).

17. Spitzer, M. Masked education? The benefits and burdens of wearing face masks in schools during the current Corona pandemic. *Trends in Neuroscience and Education* **20**, 100138 (2020).
18. Kotsia, I., Buciu, I. & Pitas, I. An analysis of facial expression recognition under partial facial image occlusion. *Image and Vision Computing* **26**, 1052–1067 (2008).
19. Neta, M. *et al.* All in the first glance: first fixation predicts individual differences in valence bias. *Cognition and Emotion* **31**, 772–780 (2017).
20. Pell, P. J. & Richards, A. Cross-emotion facial expression aftereffects. *Vision Research* **51**, 1889–1896 (2011).
21. Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J. & Kissler, J. Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS ONE* **12**, e0177239 (2017).
22. Calder, A. J. *et al.* Caricaturing facial expressions. *Cognition* **76**, 105–146 (2000).
23. Calvo, M. G., Fernández-Martín, A., Gutiérrez-García, A. & Lundqvist, D. Selective eye fixations on diagnostic face regions of dynamic emotional expressions: KDEF-dyn database. *Sci Rep* **8**, 17039 (2018).
24. Eisenbarth, H. & Alpers, G. W. Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion* **11**, 860–865 (2011).
25. Schurgin, M. W. *et al.* Eye movements during emotion recognition in faces. *Journal of Vision* **14**, 14–14 (2014).
26. Smith, M. L., Cottrell, G. W., Gosselin, F. & Schyns, P. G. Transmitting and Decoding Facial Expressions. *Psychol Sci* **16**, 184–189 (2005).
27. Bublatzky, F., Kavcıoğlu, F., Guerra, P., Doll, S. & Junghöfer, M. Contextual information resolves uncertainty about ambiguous facial emotions: Behavioral and magnetoencephalographic correlates. *NeuroImage* **215**, 116814 (2020).
28. Hassin, R. R., Aviezer, H. & Bentin, S. Inherently Ambiguous: Facial Expressions of Emotions, in Context. *Emotion Review* **5**, 60–65 (2013).
29. Kaminska, O. K. *et al.* Ambiguous at the second sight: Mixed facial expressions trigger late electrophysiological responses linked to lower social impressions. *Cogn Affect Behav Neurosci* **20**, 441–454 (2020).
30. Kinchella, J. & Guo, K. Facial Expression Ambiguity and Face Image Quality Affect Differently on Expression Interpretation Bias. *Perception* **50**, 328–342 (2021).
31. Niedenthal, P. M., Halberstadt, J. B., Margolin, J. & Innes-Ker, Åse H. Emotional state and the detection of change in facial expression of emotion. *Eur. J. Soc. Psychol.* **30**, 211–222 (2000).
32. Niedenthal, P. M., Brauer, M., Robin, L. & Innes-Ker, Å. H. Adult attachment and the perception of facial expression of emotion. *Journal of Personality and Social Psychology* **82**, 419–433 (2002).
33. Olszanowski, M., Kaminska, O. K. & Winkielman, P. Mixed matters: fluency impacts trust ratings when faces range on valence but not on motivational implications. *Cognition and Emotion* **32**, 1032–1051 (2018).

34. Sylvester, C. M., Hudziak, J. J., Gaffrey, M. S., Barch, D. M. & Luby, J. L. Stimulus-Driven Attention, Threat Bias, and Sad Bias in Youth with a History of an Anxiety Disorder or Depression. *J Abnorm Child Psychol* **44**, 219–231 (2016).
35. Ipser, A. & Cook, R. Inducing a concurrent motor load reduces categorization precision for facial expressions. *Journal of Experimental Psychology: Human Perception and Performance* **42**, 706–718 (2016).
36. Niedenthal, P. M., Mermillod, M., Maringer, M. & Hess, U. The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behav Brain Sci* **33**, 417–433 (2010).
37. Calvo, M. G., Fernández-Martín, A. & Nummenmaa, L. Perceptual, categorical, and affective processing of ambiguous smiling facial expressions. *Cognition* **125**, 373–393 (2012).
38. Elsherif, M. M., Saban, M. I. & Rotshtein, P. The perceptual saliency of fearful eyes and smiles: A signal detection study. *PLoS ONE* **12**, e0173199 (2017).
39. Grundmann, F., Epstude, K. & Scheibe, S. Face masks reduce emotion-recognition accuracy and perceived closeness. *PLOS ONE* **16**, e0249792 (2021).
40. Kret, M., Stekelenburg, J., Roelofs, K. & De Gelder, B. Perception of Face and Body Expressions Using Electromyography, Pupillometry and Gaze Measures. *Frontiers in Psychology* **4**, (2013).
41. Kret, M. E. & Fischer, A. H. Recognition of facial expressions is moderated by Islamic cues. *Cogn Emot* **32**, 623–631 (2018).
42. Lane, J. *et al.* Impacts of impaired face perception on social interactions and quality of life in age-related macular degeneration: A qualitative study and new community resources. *PLoS One* **13**, e0209218 (2018).
43. Wong, C. K. M. *et al.* Effect of facemasks on empathy and relational continuity: a randomised controlled trial in primary care. *BMC Family Practice* **14**, 200 (2013).
44. Ratcliff, R. A theory of memory retrieval. *Psychol. Rev.* **85**, 59–108 (1978).
45. Voss, A., Rothermund, K. & Voss, J. Interpreting the parameters of the diffusion model: an empirical validation. *Mem. Cognit.* **32**, 1206–20 (2004).
46. Voss, A., Rothermund, K. & Brandtstädter, J. Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology* **44**, 1048–1056 (2008).
47. Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W. & Forstmann, B. U. Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *J. Neurosci.* **32**, 2335–43 (2012).
48. Ratcliff, R. & McKoon, G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* **20**, 873–922 (2008).
49. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–74 (2007).
50. Mulder, M. J., van Maanen, L. & Forstmann, B. U. Perceptual decision neurosciences - a model-based review. *Neuroscience* **277**, 872–884 (2014).

51. Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences* **20**, 260–281 (2016).
52. Wagenmakers, E.-J. Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *Eur. J. Cogn. Psychol.* **21**, 641–671 (2009).
53. Lundqvist, D. & Litton, J. E. *The Averaged Karolinska Directed Emotional Faces - AKDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet*. (1998).
54. Guo, K. Holistic Gaze Strategy to Categorize Facial Expression of Varying Intensities. *PLoS ONE* **7**, e42585 (2012).
55. Guo, K., Soornack, Y. & Settle, R. Expression-dependent susceptibility to face distortions in processing of facial expressions of emotion. *Vision Research* **157**, 112–122 (2019).
56. Du, S. & Martinez, A. M. Wait, are you sad or angry? Large exposure time differences required for the categorization of facial expressions of emotion. *Journal of Vision* **13**, 13–13 (2013).
57. Morris, J. S., Öhman, A. & Dolan, R. J. A subcortical pathway to the right amygdala mediating “unseen” fear. *Proc. Natl. Acad. Sci. U.S.A.* <bvertical-align:super;>96</bvertical-align:super;>, 1680–1685 (1999).
58. Smith, F. W. & Schyns, P. G. Smile Through Your Fear and Sadness: Transmitting and Identifying Facial Expression Signals Over a Range of Viewing Distances. *Psychol Sci* **20**, 1202–1208 (2009).
59. Du, S. & Martinez, A. M. The resolution of facial expressions of emotion. *Journal of Vision* **11**, 24–24 (2011).
60. Dores, A. R., Barbosa, F., Queirós, C., Carvalho, I. P. & Griffiths, M. D. Recognizing Emotions through Facial Expressions: A Largescale Experimental Study. *IJERPH* **17**, 7420 (2020).
61. Wright, D. B. & Sladden, B. An own gender bias and the importance of hair in face recognition. *Acta Psychologica* **114**, 101–114 (2003).
62. Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N. & Evershed, J. K. Gorilla in our midst: An online behavioral experiment builder. *Behav Res Methods* **52**, 388–407 (2020).
63. van Ravenzwaaij, D. & Oberauer, K. How to use the diffusion model: Parameter recovery of three methods: Ez, fast-dm, and DMAT. *Journal of Mathematical Psychology* **53**, 463–473 (2009).
64. Voss, A., Nagler, M. & Lerche, V. Diffusion Models in Experimental Psychology: A Practical Introduction. *Experimental Psychology* **60**, 385–402 (2013).
65. Voss, A., Voss, J. & Lerche, V. Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Front. Psychol.* **6**, (2015).
66. Voss, A. & Voss, J. Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods* **39**, 767–775 (2007).
67. Voss, A. & Voss, J. A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology* **52**, 1–9 (2008).
68. Boehm, U. *et al.* Estimating across-trial variability parameters of the Diffusion Decision Model: Expert advice and recommendations. *Journal of Mathematical Psychology* **87**, 46–75 (2018).

69. Lerche, V. & Voss, A. Retest reliability of the parameters of the Ratcliff diffusion model. *Psychological Research* **81**, 629–652 (2017).
70. Palmer, J., Huk, A. C. & Shadlen, M. N. The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vis.* **5**, 376–404 (2005).
71. Ratcliff, R. Measuring psychometric functions with the diffusion model. *Journal of Experimental Psychology: Human Perception and Performance* **40**, 870–888 (2014).

Figures

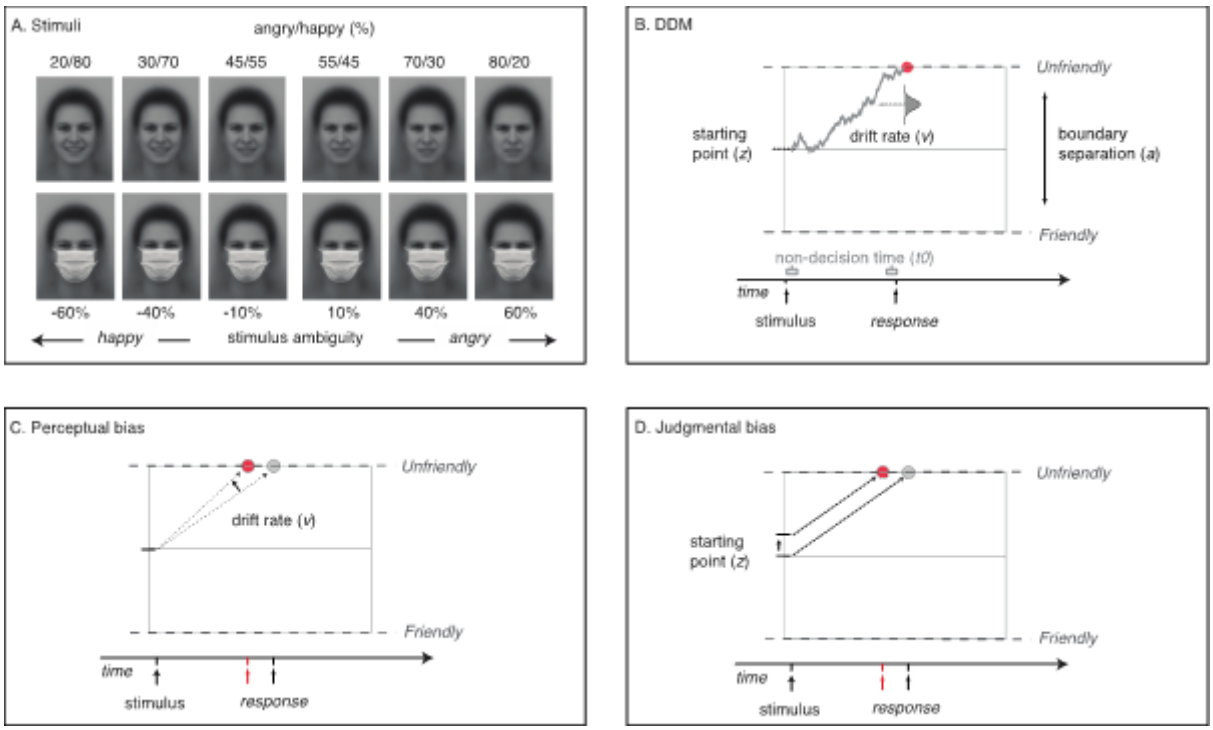

Figure 1

Ambiguous facial stimuli and model predictions. **A**. Average non-binary happy and angry faces were created by morphing the (average) male and female AKDEF[53] faces. Next, emotionally ambiguous faces were created by morphing the happy face towards the angry face in 41 incremental steps of 2.5% each. Six faces were used, each with a different angry/happy ratio. Stimulus ambiguity was defined by the difference in the terms of the ratio (Angry - Happy). **B.** The DDM (Drift Diffusion Model) represents decisions as an accumulation of noisy sensory evidence over time (drift rate *v*), which starts at starting point (*z*) and ends at a decision threshold (*a*). Non-decision time (*t0*) is the time for processes other than the decision process, such as sensory encoding and execution of the response **C**. Perceptual bias is driven by differences in information uptake at the lower perceptual level where bottom-up effects will bias the process due to the information of the mask that will be processed as evidence for the unfriendly alternative. **D**. Judgmental Bias is driven by a shift in starting point (*z*), reflecting asymmetric distances to the decision thresholds. If masked faces are associated with an unfriendly connotation, top-down effects

will bias the decision process as the unfriendly decision criterium will be closer to the starting point, resulting in faster and more unfriendly choices.

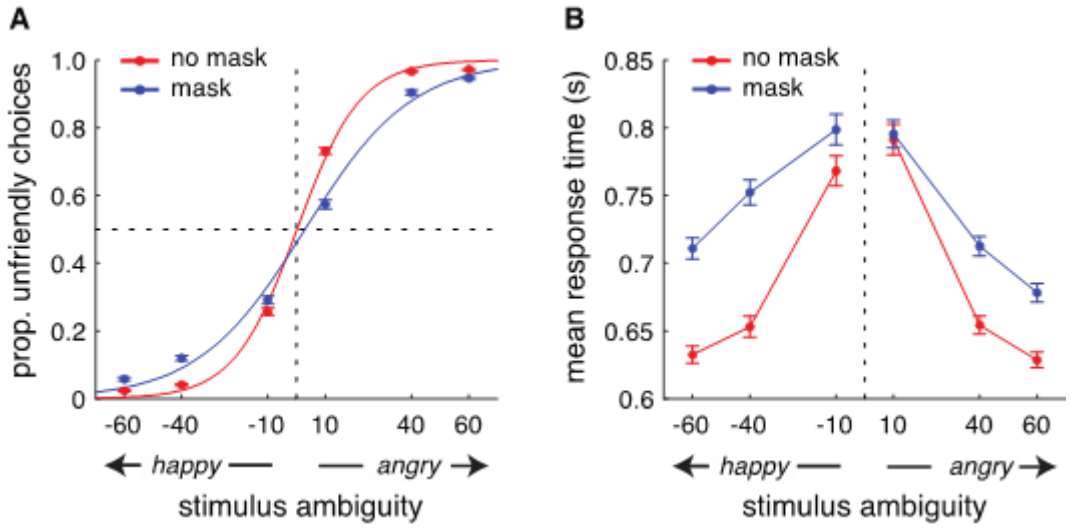

Figure 2

Descriptive data. **A**. Psychometric functions of the pooled data across participants for masked (blue) and unmasked (red) ambiguous facial expressions. The proportion of unfriendly choices is plotted as a function of stimulus intensity. **B**. Group averages of median response times (in seconds) as a function of stimulus intensity, for masked (blue) and unmasked (red) ambiguous facial expressions.

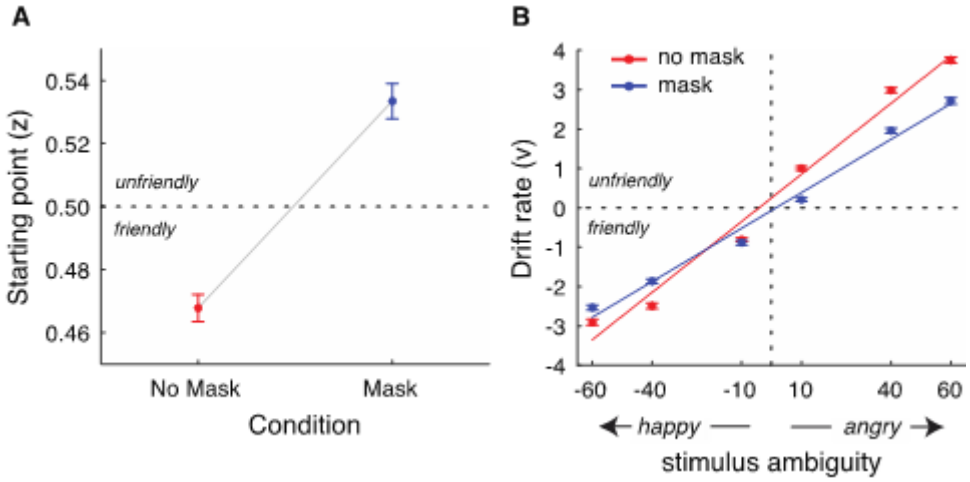

Figure 3

Judgmental and perceptual bias represented by the DDM parameters. **A**. Starting point (*z*) values for masked (blue) and unmasked (red) facial stimuli. Values larger (smaller) than 0.5 indicate a bias towards the friendly (unfriendly) decision threshold (see Fig. 1). **B**. Drift-rates as function of (signed) stimulus

ambiguity, for masked (blue) and unmasked (red) facial stimuli. A vertical shift of the regression lines (intercept) indicates a perceptual bias, with positive (negative) values for a bias towards unfriendly (friendly) sensory information.

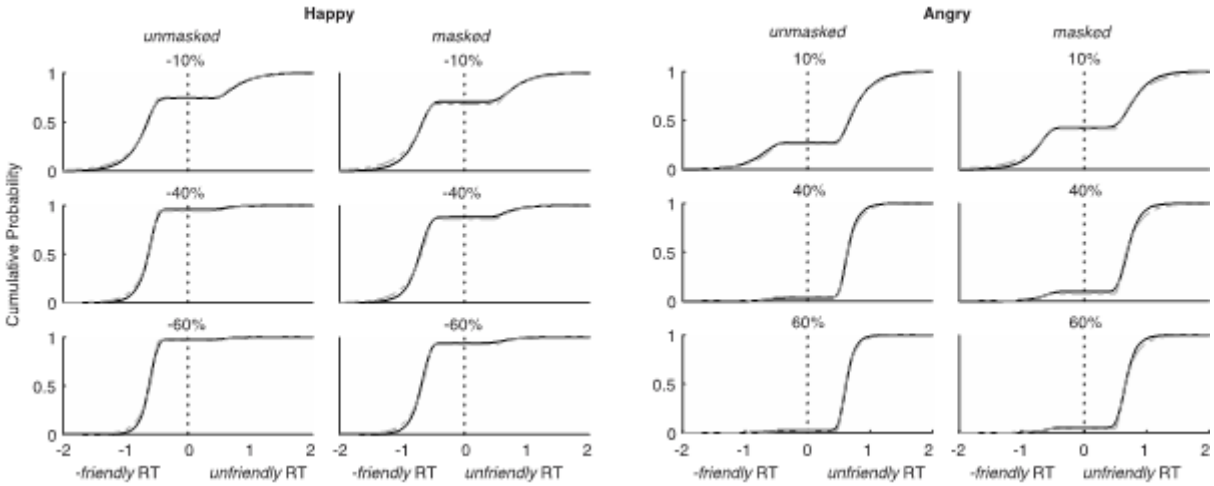

Figure 4

Goodness of fit: comparison of empirical and predicted response time distributions. The diffusion model was fitted to the conditional error and incorrect response times, collapsed across all participants (n = 136). Model fits on these group RTs distributions, resulted in 12 conditions (6 ambiguity levels for masked and unmasked faces). For each condition, the cumulative density function (CDF) was derived for friendly and unfriendly responses. Graphs represents the combined CDFs for friendly (multiplied by -1) and unfriendly responses, with the proportion of friendly responses indicated by the probability at RT = 0. The gray dotted line shows the cumulative probability of the predicted response times of the model parameters. Black line represents the empirical response times. Predicted and empirical CDFs are approximately similar, which indicates a good model fit.