Brain Behavior and Immunity xxx (xxxx) 1–17

Contents lists available at ScienceDirect

Brain Behavior and Immunity

journal homepage: www.elsevier.com/locate/ybrbi

Full-length Article

Behavioral and physiological sensitivity to natural sick faces

Tiffany S. Leung[a,*], Sarah E. Maylott[b,c], Guangyu Zeng[a], Diana N. Nascimben[a], Krisztina V. Jakobsen[d], Elizabeth A. Simpson[a]

[a] Department of Psychology, University of Miami, Coral Gables, FL, USA
[b] Department of Psychiatry & Behavioral Sciences, Duke University, Durham, NC, USA
[c] Department of Psychology, University of Utah, Salt Lake City, UT, USA
[d] Department of Psychology, James Madison University, Harrisonburg, VA, USA

ARTICLE INFO

Keywords:
Pathogen avoidance
Face perception
Sickness detection
Facial health
Visual attention
Behavioral immune system
Pupil dilation
Autonomic arousal
Physiology

ABSTRACT

The capacity to rapidly detect and avoid sick people may be adaptive. Given that faces are reliably available, as well as rapidly detected and processed, they may provide health information that influences social interaction. Prior studies used faces that were manipulated to appear sick (e.g., editing photos, inducing inflammatory response); however, responses to *naturally* sick faces remain largely unexplored. We tested whether adults detected subtle cues of genuine, acute, potentially contagious illness in face photos compared to the same individuals when healthy. We tracked illness symptoms and severity with the Sickness Questionnaire and Common Cold Questionnaire. We also checked that sick and healthy photos were matched on low-level features. We found that participants ($N = 109$) rated sick faces, compared to healthy faces, as sicker, more dangerous, and eliciting more unpleasant feelings. Participants ($N = 90$) rated sick faces as more likely to be avoided, more tired, and more negative in expression than healthy faces. In a passive-viewing eye-tracking task, participants ($N = 50$) looked longer at healthy than sick faces, especially the eye region, suggesting people may be more drawn to healthy conspecifics. When making approach-avoidance decisions, participants ($N = 112$) had greater pupil dilation to sick than healthy faces, and more pupil dilation was associated with greater avoidance, suggesting elevated arousal to threat. Across all experiments, participants' behaviors correlated with the degree of sickness, as reported by the face donors, suggesting a nuanced, fine-tuned sensitivity. Together, these findings suggest that humans may detect subtle threats of contagion from sick faces, which may facilitate illness avoidance. By better understanding how humans naturally avoid illness in conspecifics, we may identify what information is used and ultimately improve public health.

1. Introduction

Humans' sociality comes with risks. Pathogens transmitted through conspecifics (e.g., diarrheal disease, respiratory infections) have been prolific threats throughout humans' evolutionary history and continue to be global leading causes of death (Karlsson et al., 2014; Liu et al., 2012; Nabeela et al., 2014; World Health Organization, 2020). In response, humans are theorized to have evolved a behavioral immune system: a set of psychological and behavioral responses that serve as an initial defense for rapid detection, interpretation, and avoidance of sickness (Curtis et al., 2004; Schaller and Park, 2011). Consistent with this theory, people rate odors from sick people as more unpleasant and sick compared to odors from healthy people (Croy et al., 2013; Olsson et al., 2014; Regenbogen et al., 2017; Sarolidou et al., 2020a). Detecting smells of illness, however, may not be a robust method for detecting natural illness (Sarolidou et al., 2020b) and requires close contact, at which point a person may already be exposed to potentially dangerous pathogens. Sounds (e.g., coughs, sneezes) may be used from a distance (Goodwin et al., 2020; Lee et al., 2010); however, people do not appear to discriminate infectious illness from other sounds (e.g., allergies; Michalak et al., 2020). Body movements of people with experimentally-induced systemic inflammation (mimicking bacterial infection) are perceived as sicker and more tired than healthy peoples' movements (Sundelin et al., 2015). However, this cue is not available when a person is stationary. Furthermore, it is unclear to what extent *natural* sickness can also be detected from body motion.

Rapid detection and avoidance of sick people using facial information may be particularly adaptive, given that faces are more consis-

* Corresponding author.
E-mail address: tsl51@miami.edu (T.S. Leung).

https://doi.org/10.1016/j.bbi.2023.03.007
Received 16 September 2022; Received in revised form 3 March 2023; Accepted 3 March 2023
0889-1591/© 20XX

tently available than other cues and are detected and processed more rapidly than nearly any other stimulus (Crouzet et al., 2010; Jack and Schyns, 2015; Wardle et al., 2020). While cues of illness are hypothesized to be conveyed by the face (Axelsson et al., 2018; Kramer and Bressan, 2021; Regenbogen et al., 2017; Sarolidou et al., 2019), to date, it remains untested whether people can detect subtle cues of acute, contagious illness in the faces of people who are *actually* ill compared to the faces of the *same* individuals when healthy.

Prior studies have manipulated facial health. For example, endotoxin injection is an experimental model that temporarily induces systemic inflammation that makes people feel ill without being contagious (Axelsson et al., 2018; Fullerton et al., 2016; Leschak et al., 2022; Schedlowski et al., 2014). Another approach is to edit face stimuli, adding a conspicuous pathogen-cue of a facial sore (van Leeuwen and Petersen, 2018) or adjusting the location of one pupil to simulate strabismus (Tybur et al., 2020), to make people appear less healthy. While faces manipulated to appear ill, compared to unmanipulated faces, were rated as sicker (Bressan, 2021; Sarolidou et al., 2019), such studies do not answer the question of whether people can detect cues of *natural*, *genuine* acute contagious illness in faces, which may be more subtle and variable than manipulated sick faces. Therefore, there is a need to better characterize illness detection and avoidance using real cues of sickness. Furthermore, people *believe* they can visually detect illness in others (Ackerman et al., 2020), which could expose them to infectious pathogens if they fail to detect sickness. It is, therefore, important to determine whether people can actually detect sickness.

In the current study, we used face photos of people with self-reported illnesses and face photos of the same people when healthy. We hypothesized that the degree of illness would be associated with participants' behavioral and physiological responses. In three naïve samples of adults, we explored whether people differently rate (Experiment 1), visually attend to (Experiment 2), and physiologically respond to (Experiment 3) sick and healthy faces. We hypothesized that people would identify sickness and associated cues (e.g., tiredness) from faces (Experiment 1), visually discriminate sick from healthy faces when health is task-irrelevant (Experiment 2), and differ in pupil dilation to sick and healthy faces (Experiment 3).

2. Experiment 1

When people are sick, they may exhibit lassitude—an emotion associated with sickness, characterized by slack facial muscles, drooping eyelids, and slightly parted lips (Schrock et al., 2020). Sick people may also appear tired (Axelsson et al., 2018) and display negative emotions (e.g., sadness; Sarolidou et al., 2019). Lassitude expressions are theorized to elicit care from others; however, such care requires close contact, which comes with costs, including increasing the helper's infection risk (Tybur et al., 2013). The capacity to identify genuine cues of infectious sickness in others may, therefore, be valuable in helping people to navigate approach-avoidance motivations.

In one study that used faces from men with HIV or herpes and a healthy control group, participants categorized faces as sick or healthy at levels above chance (Tskhay et al., 2016). These findings are consistent with the behavioral immune system hypothesis, suggesting people may be sensitive to genuine sickness. However, in this study the sick and healthy faces were from different individuals, leaving open the possibility of other differences among these groups (e.g., socioeconomic status), making it difficult to know the extent to which participants relied on cues of illness, specifically. Further, it remains untested whether people are sensitive to facial cues of more infectious diseases, such as those transmitted through the air during brief encounters (e.g., COVID-19).

To address this gap, in Experiment 1 we examined whether people differently rated actual sick and healthy faces from the same individuals. We included faces with illnesses that are transmissible through shared spaces or brief contact, and therefore, would be evolutionarily adaptive to avoid. In line with previous reports using experimentally manipulated sick faces (Axelsson et al., 2018; Sarolidou et al., 2019), we predicted that natural sick faces would be rated as more sick, dangerous, avoided, unpleasant, tired, and negative compared to healthy faces (Prediction 1a). We also predicted that these ratings would be negatively associated with donor-reported health (Prediction 1b) and would be positively associated with one another (Prediction 1c).

2.1. Materials and methods

2.1.1. Participants

Adult participants (N = 199) were recruited from a research participant pool to rate the faces (see Table 1 for demographic details). Raters received course credit for their participation. The University of Miami Institutional Review Board approved this study.

2.1.2. Stimuli

We used 32 neutral/relaxed face photos collected from 16 individuals (face photo donors), once when sick and once when healthy. Sick faces were donated by people who were actually sick. Donors were permitted to submit photos of their sick faces at various points of illness (e.g., initial onset, peak), which allowed us to capture variability in symptom severity across faces while donors were still contagious. Healthy faces were from the same individuals prior to illness or once they were fully recovered. Sickness and recovery were confirmed through self-report (for full survey, see Supplementary Materials). Sick and healthy photos were taken an average of 9.30 months apart. Face photo donors consented to sharing their photos for use in future studies and to be included in a face database and scientific publications (e.g., journal articles, conference presentations). Face photo donors were instructed to take a straight-on photo at eye-level (as opposed to from above/below or from the side), while making eye-contact with the camera and maintaining a neutral and relaxed facial expression, in a well-lit location. When taking their second face photo, donors were instructed to take the photo in the same location and at the same time of day as their first photo and to replicate the hairstyle and lighting as much as possible. We also asked donors to avoid make-up, eye glasses, and jewelry. However, 2 of the 16 donors were wearing glasses in one photo, so we ensured that they were wearing them in the other one as well (in both the sick and healthy photos), for consistency. Donors shared the original, uncropped photos, which we then standardized using Adobe

Table 1
Participant demographics.

	Experiment 1 (N = 199)	Experiment 2 (N = 50)	Experiment 3 (N = 112)
Age	18–29 years, M = 19.4, SD = 1.5	17–22 years, M = 18.48, SD = 0.84	16–29 years, M = 19.39, SD = 1.80
Gender			
Woman	106 (53%)	36 (72%)	80 (71%)
Man	93 (47%)	14 (28%)	32 (29%)
Race			
American Indian or Alaskan Native	1 (1%)	0 (0%)	0 (0%)
Asian	26 (13%)	1 (2%)	9 (8%)
Black or African American	25 (13%)	4 (8%)	17 (15%)
Middle Eastern	3 (2%)	1 (2%)	3 (3%)
Native Hawaiian or Other Pacific Islander	1 (1%)	0 (0%)	0 (0%)
White	123 (62%)	39 (78%)	75 (67%)
Multiracial	15 (8%)	3 (6%)	8 (7%)
Unknown/Other	5 (3%)	2 (4%)	0 (0%)
Ethnicity			
Hispanic or Latino	41 (21%)	14 (28%)	27 (23%)
Not Hispanic or Latino	152 (76%)	35 (70%)	85 (76%)
Other	6 (3%)	1 (2%)	0 (0%)

Photoshop. Photos were sized using inter-pupil distance, cropped, and given uniform white clothing and backgrounds (for examples, see Fig. 1A). We used the Saliency Toolbox to confirm that features related to low-level salience (e.g., color, luminance, contour) did not differ be-

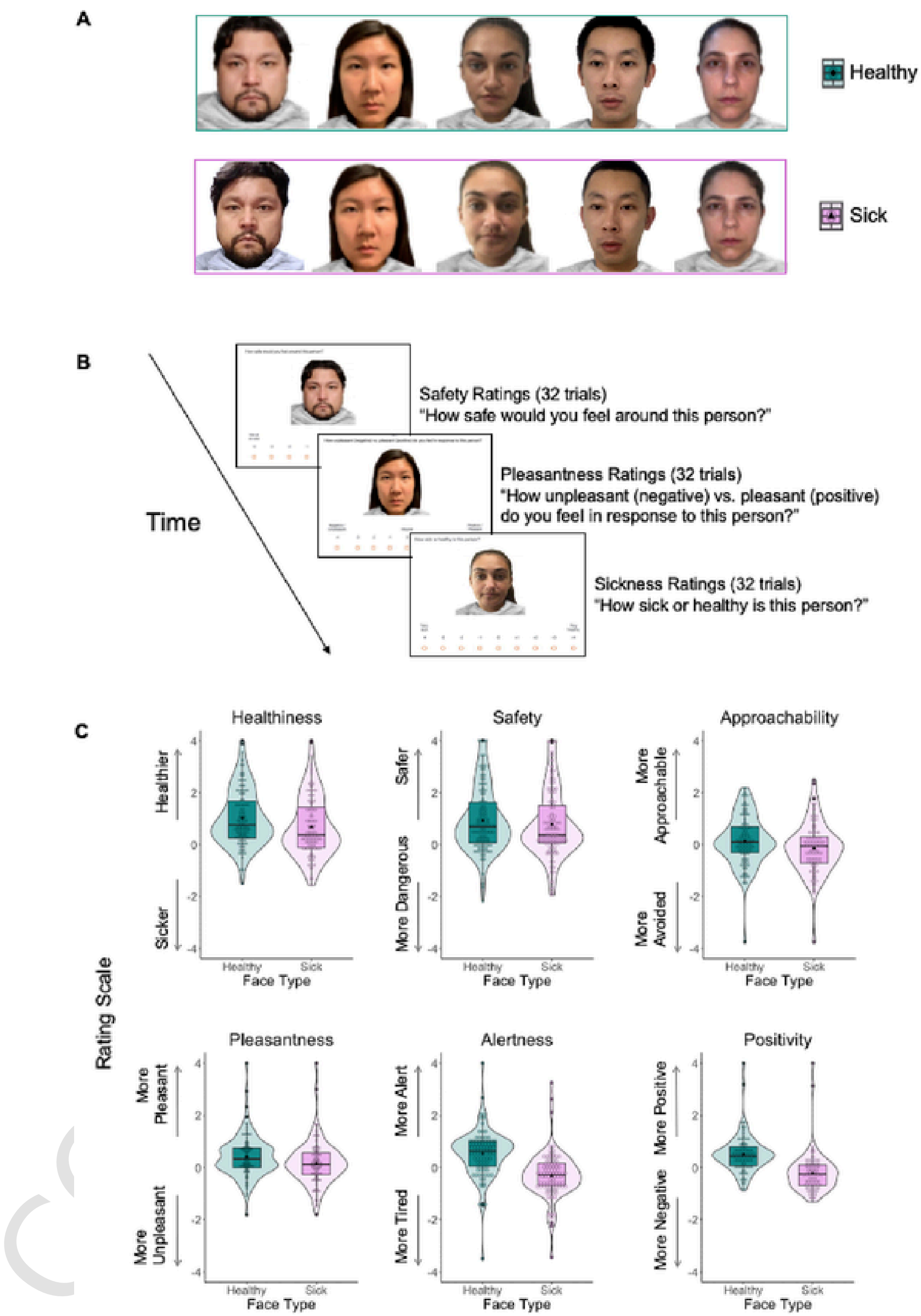

Fig. 1. (A) Healthy (top row; teal blue) and sick (bottom row; pink) face stimuli examples. (B) Example study procedure (Version 1): raters completed 32 ratings from −4 to + 4 in each dimension before moving to the next rating dimension. (C) Healthy faces (teal blue) were rated as healthier (less sick), safer (less dangerous), more approachable (less likely to be avoided), eliciting more pleasant feelings, more awake (less tired), and more positive in their expressions compared to sick faces (pink). In contrast, sick faces were rated as sicker (less healthy), more dangerous (less safe), less approachable (more likely to be avoided), eliciting more unpleasant feelings, less awake (more tired), and negative in their expressions compared to healthy faces. Horizontal lines within the boxplots indicate the medians. Black dots indicate means. Hinges of the boxplots show the first (bottom) and third (top) quartiles. The whiskers extend up to 1.5 × Interquartile Range (IQR; distance between top and bottom hinges), above and below the hinges. Colored dots show binned raw data (e.g., each individual participant's rating). Violin plots show the frequency distributions of the ratings. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tween sick and healthy faces (see Supplementary Method; Ho-Phuoc et al., 2010; Walther and Koch, 2006).

To confirm that individuals were sick and healthy in each of their photos, we adapted questions from validated measures of health from national public health surveys and research, including symptoms (e.g., headache, cough, nausea), severity, and diagnoses (e.g., COVID-19; Andreasson et al., 2018; Centers for Disease Control and Prevention; 2021; Powell et al., 2008; see Supplementary Materials for full list of survey questions), collected through a Qualtrics survey (qualtrics.com). We confirmed that donors rated themselves as being in poorer health when they donated the sick face photos than when they donated the healthy face photos (Table 2). Donors had more severe symptoms (e.g., fever) when sick, reflected in a higher Sickness Questionnaire (SicknessQ; Andreasson et al., 2018) total score when they were sick ($M = 16.31$, $SD = 6.55$) compared to when they were healthy ($M = 1.36$, $SD = 2.67$), $t(12) = 8.31$, $p < 0.001$, $d = 2.31$, and a higher Common Cold Questionnaire (CCQ; Powell et al., 2008) total score when they were sick ($M = 15.62$, $SD = 5.42$) compared to when they were healthy ($M = 1.57$, $SD = 3.57$), $t(12) = 8.42$, $p < 0.001$, $d = 2.33$. For example, when donors were sick, 92% reported coughing, nasal discharge, nasal obstruction, sneezing, and/or sore throat, and 62% reported abdominal pain, diarrhea, nausea, and/or vomiting (vs. only 14% and 7%, respectively, when healthy). In addition, when sick, 85% of donors were unable to work, perform daily activities, and/or reported that they got care from a doctor or other healthcare professional (vs. only 7% when healthy). For a full list of specific illnesses, see Table S1. CCQ and SicknessQ scores positively correlated with one another, $r(25) = 0.90$, $p < 0.001$, confirming that they were both capturing the degree of illness/health.

2.1.3. Ratings

Each image was rated on 6 dimensions (Axelsson et al., 2018; Campbell et al., 2021; Sarolidou et al., 2019) using 9-point scales: (1) “How sick or healthy is this person?” on a scale from “very sick” (−4) to “very healthy” (+4); (2) “How safe would you feel around this person?” on a scale from “not at all safe” (−4) to “very safe” (+4); (3) “How strong is your urge to approach or avoid this person?” on a scale from “avoid” (−4) to “approach” (+4); (4) “How unpleasant (negative) vs. pleasant (positive) do you feel in response to this person?” on a scale from “negative/unpleasant” (−4) to “positive/pleasant” (+4) with the midpoint labeled “neutral” (0); (5) “How tired (sleepy) vs alert (awake) is this person?” on a scale from “very tired” (−4) to “very alert” (+4); (6) “How is this person feeling?” on a scale from “very negative” (−4) to “very positive” (+4) with the midpoint labeled “neutral” (0). For each rating, the photo was visible along with the question and response options. Out of 13,728 trials, 123 were lost due to experimenter error (<1%).

Table 2
Experiment 1 donor-reported symptoms and severity for healthy and sick face photos.

	Healthy Faces	Sick Faces				
	M (*SD*)	*M* (*SD*)	*t*	*df*	*p*	*d*
Chills	0.00 (0.00)	1.23 (1.01)	4.38	12	<0.001*	1.22
Headache	0.29 (0.83)	2.00 (1.16)	4.88	12	<0.001*	1.35
Joint pains	0.14 (0.36)	1.08 (1.12)	3.49	12	0.004#	0.97
Muscle or body aches	0.14 (0.36)	1.92 (0.95)	6.30	12	<0.001*	1.75
Fatigue or tiredness	0.36 (0.63)	2.46 (0.78)	8.68	12	<0.001*	2.41
Nausea	0.14 (0.54)	1.00 (0.91)	3.40	12	0.005#	0.94
Vomiting	0.00 (0.00)	0.15 (0.38)	1.48	12	0.165	0.41
Diarrhea	0.00 (0.00)	0.08 (0.28)	1.00	12	0.337	0.28
Abdominal pain	0.00 (0.00)	0.62 (0.87)	2.55	12	0.025#	0.71
Sneezing	0.14 (0.36)	1.00 (1.08)	2.86	12	0.014#	0.79
Sore throat	0.00 (0.00)	1.38 (1.04)	4.78	12	<0.001*	1.33
Nasal discharge	0.21 (0.58)	0.85 (0.80)	2.31	12	0.040#	0.64
Nasal obstruction	0.07 (0.27)	0.85 (0.99)	2.74	12	0.018#	0.76
Cough	0.07 (0.27)	1.00 (1.08)	2.80	12	0.016#	0.78

Note. Face donors reported significantly better health (i.e., less severe sickness symptoms) at the time of donating their healthy face photo (left) compared to the time they donated their sick face photo (right). For each symptom, donors reported None (0), Mild (1), Moderate (2), or Severe (3). Health rating data from 3 donors were unavailable, as reported above. Means (*M*) and standard deviations (*SD*) are reported for all available symptom scores. Degrees of freedom (*df*) and Cohen's *d* effect sizes are reported for each paired samples *t* test, #$ps < 0.05$, *$ps < 0.0036$ (Bonferroni correction: $\alpha_{crit} = 0.05/14$).

2.1.4. Procedure

Ratings were completed online through Qualtrics (qualtrics.com) between April 2021 and December 2022. Raters completed the study on a computer ($n = 135$), smartphone ($n = 5$), or tablet ($n = 3$). Raters rated all 32 images on one dimension at a time in a block of 32 trials before moving on to the next rating blocks. Within each block, images were presented in a random order. To avoid fatigue effects, raters were randomly assigned to complete only 3 ratings (3 blocks) for the 32 images instead of all 6 ratings, for 96 trials total per person. Through random assignment, approximately half the raters ($N = 109$) rated safety, pleasantness, and healthiness (Version 1; Fig. 1B). We asked participants to rate healthiness after they completed the other ratings to avoid priming their attention to health-related cues. The other half of the raters ($N = 90$) rated approachability, tiredness, and the positivity of the face (Version 2). At the end of the study, raters reported demographic information, and whether they recognized any of the people in the photos. To eliminate the potential effect of familiarity, data from six raters were excluded because they reported recognizing one or more faces. The survey took approximately 20 min.

2.1.5. Analysis

We first confirmed that our donors had more symptoms of illness when they contributed their sick face photo than when they contributed their healthy face photo by conducting paired samples *t* tests using two measures of overall health: the CCQ total score (Powell et al., 2008) and SicknessQ total score (Andreasson et al., 2018). Donors rated 14 CCQ symptoms (e.g., “fatigue or tiredness”) on a 4-point scale (none = 0, mild = 1, moderate = 2, severe = 3), and total scores were calculated by summing symptom severities. Similarly, they rated 10 SicknessQ statements (e.g., “I didn't wish to do anything at all”) on a 4-point scale (disagree = 0, agree somewhat = 1, mostly agree = 2, agree = 3), and total scores were calculated by summing statement scores. For both total scores, higher scores indicate overall more severe symptoms and/or a greater number of symptoms.

All analyses were conducted in R (version 4.2.1). For our primary analysis, we conducted a linear mixed effects model using the “lme4” (Bates et al., 2015) and “lmerTest” (Kuznetsova et al., 2017), including a fixed effect for health condition and random effects of participants for health condition, to compare participants' sick face ratings to their healthy face ratings for each of the 6 ratings (healthiness, safety, approachability, pleasantness, alertness, and positivity), with Bonferroni corrections ($\alpha_{crit} = 0.05/6 = 0.0083$) (Prediction 1a). We also conducted 6 multilevel correlations, controlling for random effects of participants, using the “correlation” package (Makowski et al., 2022) for each of the 6 ratings to check whether, at the trial level, the 6 face ratings were related to donor-reported health in the CCQ and SicknessQ, with Bonferroni corrections ($\alpha_{crit} = 0.05/6 = 0.0083$) (Prediction 1b). Given that our limited sample size of face donors ($N = 16$) might lead to a lack of power and an increased estimation bias for the donor-level random effects (Maas and Hox, 2005; Westfall et al., 2014), we only controlled for random effects of the participants for our primary analysis and multilevel correlations. Finally, we checked whether, at the stimulus level (32 face photos), the 6 ratings were correlated with one another with 15 Pearson correlations, with Bonferroni corrections (α_{crit} = 0.05/15 = 0.0033) (Prediction 1c).

2.2. Experiment 1 results

We found that sick faces, compared to healthy faces, were rated as sicker (less healthy), more dangerous (less safe), more likely to be avoided (less approachable), more unpleasant to view (less pleasant), more tired (less awake), and exhibiting more negative affect (less positive affect; Fig. 1C; Table 3; Prediction 1a). Further, we found that face ratings negatively correlated with donor-reported health (Table 4; Prediction 1b), suggesting that poorer reported health was associated with lower ratings of faces' healthiness, safety, approachability, pleasantness, alertness, and positivity. We also found that the ratings were positively associated with one another (Fig. 2; Table 5; Prediction 1c), suggesting that these dimensions may be capturing distinct but related underlying facial qualities.

Table 3
Experiment 1 healthy and sick face ratings.

	Healthy Faces	Sick Faces	Fixed Effects				
	M (*SD*)	*M* (*SD*)	*b*	*t*	*df*	*p*	*d*
Healthiness	1.11 (1.97)	0.78 (2.12)	0.32	5.37	1721	< 0.001*	0.26
Safety	1.02 (2.18)	0.83 (2.20)	0.19	3.06	2753	0.002*	0.12
Approachability	0.13 (1.91)	-0.17 (1.89)	0.29	4.51	2651	< 0.001*	0.18
Pleasantness	0.42 (1.90)	0.17 (1.92)	0.24	4.06	3173	< 0.001*	0.14
Alertness	0.54 (2.08)	-0.31 (2.18)	0.85	11.51	2513	< 0.001*	0.46
Positivity	0.50 (1.82)	-0.22 (1.96)	0.73	10.88	1868	< 0.001*	0.50

Note. Sick and healthy face ratings on 6 dimensions on a 9-point Likert scale: very sick (−4) to very healthy (+4), not at all safe (−4) to safe (+4), avoid (−4) to approach (+4), unpleasant/negative (−4) to pleasant/positive (+4) in how the participant feels, and whether the face appears very tired (−4) to very alert (+4), and has an expression that is very negative (−4) to very positive (+4). Means (*M*), standard deviations (*SD*), degrees of freedom (*df*), and Cohen's *d* effect sizes calculated from the linear mixed effects models are reported for each rating, **ps* < 0.0083 (Bonferroni corrections: α_{crit} = 0.05/6). Degrees of freedom (*df*) varied across correlations due to missing data (162 missing trials total; less than 1% of the data).

Table 4
Multilevel correlations between face ratings and donor-reported health, controlling for random effects of participants, in Experiment 1.

Face Rating	CCQ		SicknessQ	
	r (df)	*p*	*r (df)*	*p*
Healthiness	-0.11 (2905)	< 0.001*	-0.10 (2905)	< 0.001*
Safety	-0.09 (2904)	< 0.001*	-0.05 (2904)	0.005*
Approachability	-0.12 (2372)	< 0.001*	-0.11 (2372)	< 0.001*
Pleasantness	-0.10 (2862)	< 0.001*	-0.06 (2862)	0.002*
Alertness	-0.21 (2411)	< 0.001*	-0.18 (2411)	< 0.001*
Positivity	-0.24 (2412)	< 0.001*	-0.22 (2412)	< 0.001*

Note. Negative correlations indicate that poorer health reported by face donors–i.e., higher scores in the Common Cold Questionnaire (CCQ) and in the Sickness Questionnaire (SicknessQ)–were associated with lower face rating scores (lower healthiness, safety, approachability, pleasantness, alertness, and positivity). These correlations were conducted at the trial level to capture each participant's rating of each individual face to determine if those ratings are associated with donor-reported health. *Indicates statistical significance with Bonferroni correction (*ps* < 0.0083). Degrees of freedom (*df*) varied across correlations due to missing data (84 missing trials total; less than 1% of the data).

2.3. Experiment 1 discussion

As predicted, participants rated sick faces as sicker and more likely to be avoided than healthy faces from the same individuals (Prediction 1a), consistent with theories that humans evolved behavioral adaptations to lessen contact with pathogens (Schaller and Park, 2011; Kramer and Bressan, 2021). These findings—that illness can be identified from the face and avoided—add to a growing body of evidence that humans are sensitive to potential cues of illness, including smell (Moshkin et al., 2012; Shirasu and Touhara, 2011) and biological motion (Lasselin et al., 2020; Sundelin et al., 2015). Our findings are also in line with a report that people can identify men infected with viruses (HIV/herpes) from a different group of virus-negative men using only pictures of faces (Tskhay et al., 2016). The current study extends these findings to suggest that healthy people appear different from sick people, *even when the photo is of the same person*, and even when the illness is acute rather than chronic. Together, these studies suggest that sickness is readily detectable from faces.

The present study is the first—to our knowledge—to use *actual* sick and healthy face photos from the same individuals. In contrast to prior studies that induced immune responses in face donors and edited face images to simulate illness (Axelsson et al., 2018; Fullerton et al., 2016; van Leeuwen and Petersen, 2018; Schedlowski et al., 2014), our face photo donors had genuine, subtle illness cues, reflective of the information available in real-world encounters. Given that our face photo donors experienced varying symptom severities and types, some resulting from airborne diseases (e.g., COVID-19), our ratings allowed us to capture responses to cues of pathogens transmissible through brief social interaction. Some of our face photo donors experienced symptoms of non-contagious disorders (see Table S1); however, we lacked power in the current study to explore specific sickness subtypes. Future studies are needed to compare perception of infectious and non-infectious sick faces to determine whether the latter elicit greater avoidance than the former. However, our inclusion of a broad variety of symptoms highlights humans' broad ability to detect features of sickness that may be present during both infectious and non-infectious illness. We found that, even with these more naturalistic stimuli, people were still sensitive to cues indicating sickness. By using sick and healthy face images from the same donors, our study was able to rule out alternative factors that vary among individuals. Together, our findings highlight that humans may have an adaptation to detect and avoid signs of contagion, as reflected in people's faces.

We also found that donor-reported health was associated with face ratings: healthier face photos made participants feel safer and more pleasant and were rated as being more alert and positive (Prediction 1b). Participants' ratings were also correlated with each other, perhaps reflecting that sicker faces were perceived as being less safe, approachable, pleasant, alert, and positive (Prediction 1c). These findings are consistent with studies using experimentally simulated sick faces (i.e., induced-immune responses; Axelsson et al., 2018; Leschak et al., 2022; Sarolidou et al., 2019). For example, people report that they would feel safer around people with healthy faces compared to faces of people simulated to be sick (Leschak et al., 2022). In addition, our findings are consistent with prior research that the perception of facial health and facial emotions are linked. For example, happy faces—regardless of their actual health—are rated as healthier than neutral faces (Jones et al., 2018). Here, we systematically explored facial health of naturally sick people and found that it was associated with the perception of emotions in the face, a novel finding. Together, these findings suggest robust associations among these interpersonal dimensions—in both simulated and natural illness—and likely, in combination, may affect how people perceive others.

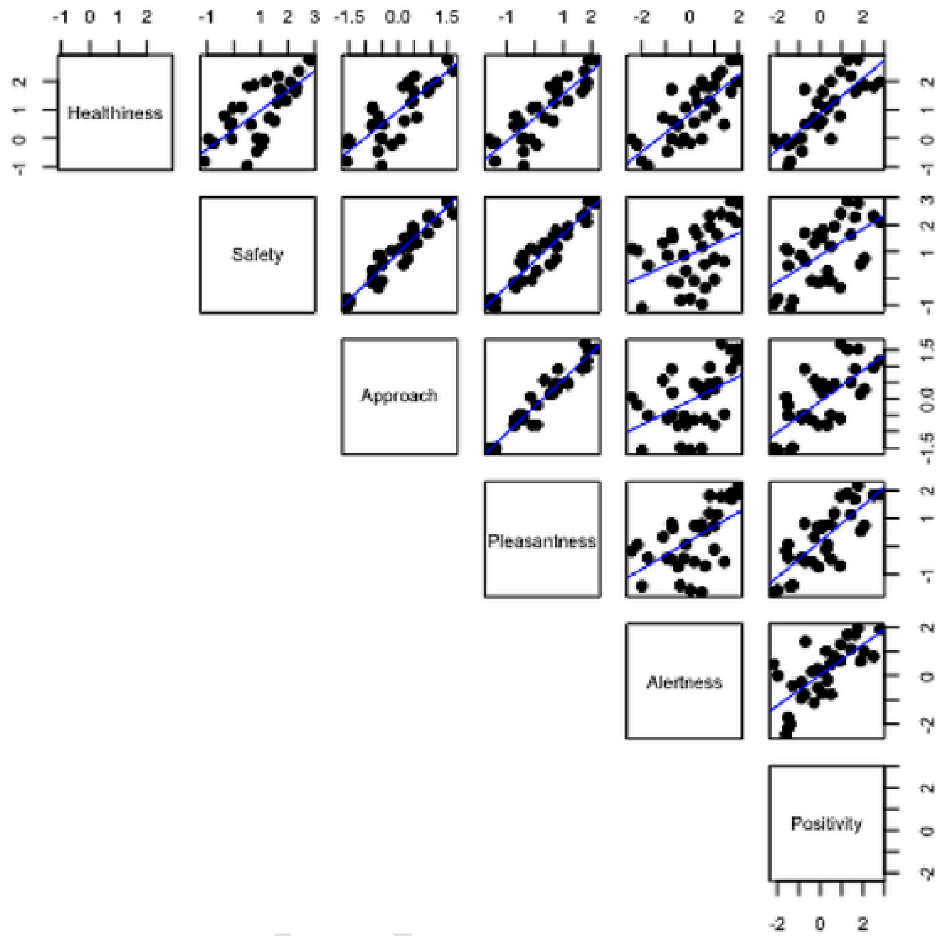

Fig. 2. Pearson correlations among face ratings in Experiment 1. Dots represent average ratings across participants for each face donor. Regression lines are depicted in blue. Face ratings positively correlated with one another (see Table 5 for p-values). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
Correlations among face ratings in Experiment 1.

	Safety	Approach	Pleasantness	Alertness	Positivity
Healthiness	$r(32) = 0.73$, $p < 0.001$*	$r(32) = 0.82$, $p < 0.001$*	$r(32) = 0.86$, $p < 0.001$*	$r(32) = 0.77$, $p < 0.001$*	$r(32) = 0.81$, $p < 0.001$*
Safety		$r(32) = 0.95$, $p < 0.001$*	$r(32) = 0.95$, $p < 0.001$*	$r(32) = 0.43$, $p = 0.014$#	$r(32) = 0.59$, $p < 0.001$*
Approach			$r(32) = 0.96$, $p < 0.001$*	$r(32) = 0.48$, $p = 0.005$*	$r(32) = 0.69$, $p < 0.001$*
Pleasantness				$r(32) = 0.56$, $p < 0.001$*	$r(32) = 0.79$, $p < 0.001$*
Alertness					$r(32) = 0.71$, $p < 0.001$*

Note. *$ps < 0.0033$ (Bonferroni correction: $\alpha_{crit} = 0.05/15$), #$ps < 0.05$.

3. Experiment 2

While Experiment 1 revealed that humans can explicitly identify genuine facial cues of sickness, including contagious disease, the extent to which people do this automatically, even when health is task-irrelevant (i.e., not instructed to attend to health-relevant cues), remains largely unexplored. We hypothesized that people may identify subtle cues of health from faces, consistent with the theorized behavioral immune system (Curtis, 2014; Schaller and Park, 2011). Attention is more efficiently captured by threats—including evolutionarily-relevant threats (e.g., snakes; Shapouri and Martin, 2021) and social threats (e.g., angry faces; Berdica et al., 2018; Feldmann-Wüstefeld et al., 2011)—relative to non-threats. We, therefore, predicted that sick faces—which are theorized to be evolutionarily-relevant—may be perceived as threatening, and therefore, may capture attention more rapidly (Prediction 2a) and hold attention for longer (Prediction 2b), compared to healthy faces. We also predicted that face viewing times would be associated with donor-reported health (Prediction 2c).

Viewers may be sensitive to expressions of lassitude, and therefore may be differentially attentive to the eye and mouth regions of sick and healthy faces. For example, people look longer to the eyes of faces generated by Artificial Intelligence when they are labeled as healthy than when they are labeled as sick, but look longer to the mouth region when faces are labeled as sick than when they are labeled as healthy (Federico et al., 2021). Further, experimentally-induced sickness is associated with viewers reporting redder eyes, paler lips, and droopier mouths (Axelsson et al., 2018), consistent with the proposal that the eye and mouth regions convey health information. We, therefore, predicted that participants would look at sick eyes and mouths more rapidly (Prediction 2d) and for longer durations (Prediction 2e) compared to healthy eyes and mouths, but would not show these differences for other face regions (e.g., noses). We also predicted that face region viewing times would be associated with donor-reported health (Prediction 2f).

Additionally, because the face photos used in Experiment 1 contained external features (e.g., hair, ears), we removed this information by cropping the photos (for examples: Fig. 3A) to test whether the inner facial features more specifically can be used. We wanted to ensure that any differential looking between the sick and healthy faces was due to qualities of the faces themselves, rather than external features (e.g., more styled hair when healthy compared to sick). We displayed the faces in pairs, presented side-by-side–one sick and one healthy, both from the same person–to enable us to track viewing patterns in how observers' attention is captured and held by each face relative to the other. Exploring people's natural viewing patterns of faces may reveal what face regions are useful for distinguishing sick and healthy faces.

3.1. Methods

3.1.1. Participants

We recruited a new set of participants ($N = 50$), who did not participate in Experiment 1 (see Table 1 for demographic details). An a priori power analysis in G*Power (Erdfelder et al., 1996) determined that a sample size of 41 participants would provide 80% power to detect small-moderate effect sizes ($d = 0.40$) for our analyses comparing participants' looking patterns for sick faces and healthy faces (paired samples *t* tests).

Participants received course credit for their participation. The University of Miami Institutional Review Board approved this study.

3.1.2. Stimuli

We used 44 neutral/relaxed face photos collected from 22 individuals, once when sick and once when healthy. These included the 32 images from Experiment 1, as well as an additional 12 images from 6 individuals (3 male donors, 3 female donors; 6 White; 2 Hispanic or Latino, 4 Not Hispanic or Latino). Among these 6 individuals, none were wearing glasses. As in Experiment 1, we obtained consent to use face photos and collected health information from face donors for Experiment 2 (Table 6). Sick and healthy photos were taken an average of 7.29 months apart. We again confirmed that donors had more severe symptoms (e.g., fever) when sick, reflected in a higher SicknessQ (Andreasson et al., 2018) total score when they were sick ($M = 17.06$, $SD = 6.46$) compared to when they were healthy ($M = 1.42$, $SD = 2.29$), $t(17) = 10.36$, $p < 0.001$, $d = 2.44$, and a higher CCQ (Powell et al., 2008) total score when they were sick ($M = 15.78$, $SD = 5.17$) compared to when they were healthy ($M = 1.42$, $SD = 3.06$), $t(17) = 10.67$, $p < 0.001$, $d = 2.52$. For example, when donors were sick, 100% reported experiencing aches, soreness, joint pains, and/or headaches, and 100% reported fatigue and/or tiredness. Additionally, 83% reported coughing, nasal discharge, nasal obstruction, sneezing, and/or sore throat, and 72% reported abdominal pain, diarrhea, nausea, and/or vomiting. Finally, 83% of donors reported chills and/or fever, and 83% were unable to work, perform daily activities, and/or reported that they got care from a doctor or other healthcare professional. CCQ and SicknessQ scores positively correlated with one another, $r(35) = 0.91$, $p < 0.001$, confirming that they were both capturing the degree of illness/health. We used identical crops on the sick and healthy faces, and again ensured they were sized matched on inter-pupil distance.

We displayed faces in pairs, with the same person's face shown twice, one image of the person's face when they were sick and a second

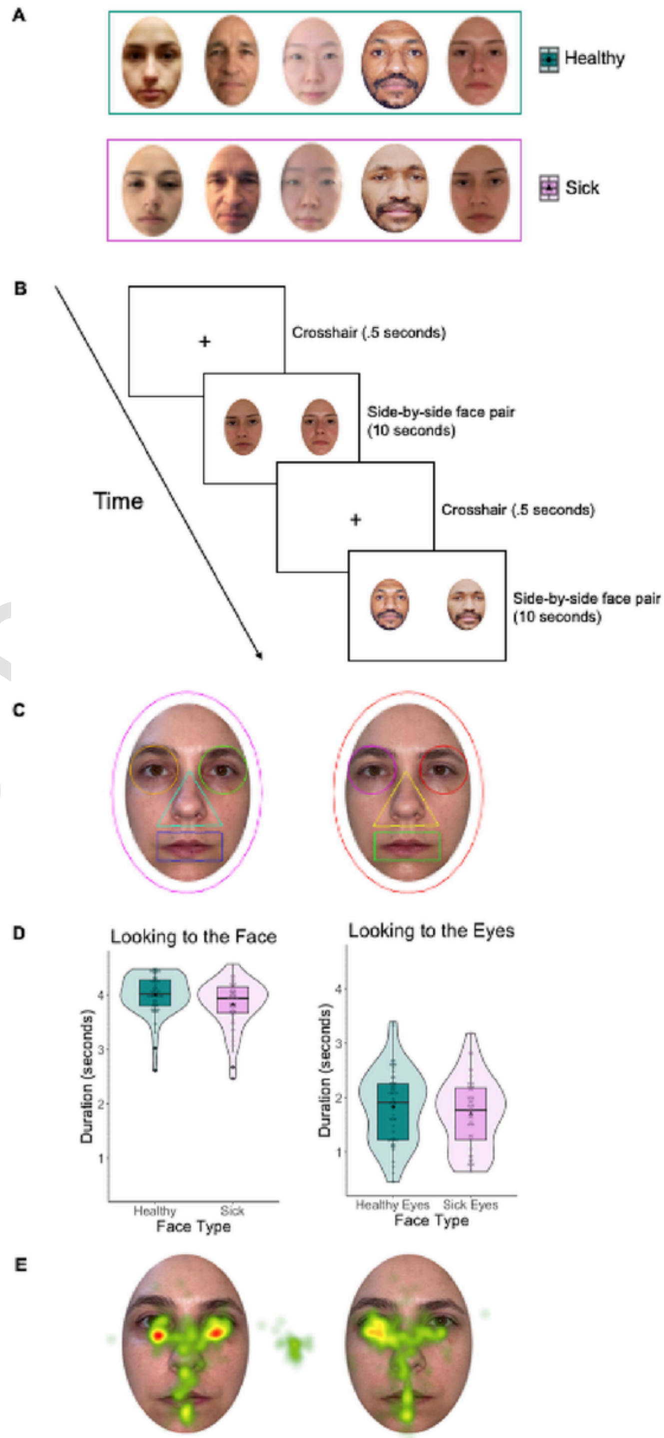

Fig. 3. (A) Healthy (top row; teal blue) and sick (bottom row; pink) oval-cropped face stimuli examples. (B) Example of free-viewing study procedure. Participants were free to look wherever they wanted to look. (C) Example of areas of interest drawn around the eyes, nose, and mouth of a sick (left) and healthy (right) face. (D) Look duration (seconds) to the faces (left graph) and eyes (right graph) for the healthy faces (teal blue, circles) and sick faces (pink, triangles). Horizontal lines within the boxplots indicate the medians. Black circles/triangles indicate means. Hinges of the boxplots show the first (bottom) and third (top) quartiles. The whiskers extend up to 1.5 × Interquartile Range, above and below the hinges. Dots are binned raw data (e.g., each individual participant's look duration). Violin plots show the frequency distributions of the look durations. (E) Heat map example of time looking at a healthy face (left)

and a sick (right) face pair, averaged across all participants who viewed this image pair. Red indicates longer fixation durations, yellow indicates moderate looking, green indicates a little looking, and regions without color indicate no looking. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Experiment 2 donor-reported symptoms and severity for healthy and sick face photos.

	Healthy Faces	Sick Faces				
	***M* (*SD*)**	*M* (*SD*)	*t*	*df*	*p*	*d*
Chills	0.00 (0.00)	1.17 (1.04)	4.5	17	< 0.001*	1.12
Headache	0.26 (0.73)	1.94 (1.16)	5.72	17	< 0.001*	1.35
Joint pains	0.11 (0.32)	1.12 (1.17)	3.89	16	0.001*	0.94
Muscle or body aches	0.11 (0.32)	1.83 (0.92)	7.63	17	< 0.001*	1.80
Fatigue or tiredness	0.42 (0.61)	2.56 (0.71)	10.76	17	< 0.001*	2.54
Nausea	0.11 (0.46)	1.17 (0.99)	4.49	17	< 0.001*	1.06
Vomiting	0.00 (0.00)	0.28 (0.75)	1.57	17	0.135	0.37
Diarrhea	0.00 (0.00)	0.39 (0.98)	1.69	17	0.110	0.40
Abdominal pain	0.00 (0.00)	0.83 (0.99)	3.59	17	0.002*	0.85
Sneezing	0.11 (0.32)	0.94 (1.06)	3.39	17	0.003*	0.80
Sore throat	0.00 (0.00)	1.33 (1.09)	5.22	17	< 0.001*	1.23
Nasal discharge	0.21 (0.54)	0.67 (0.77)	2.05	17	0.057	0.48
Nasal obstruction	0.05 (0.23)	0.67 (0.91)	2.83	17	0.012#	0.67
Cough	0.05 (0.23)	0.94 (1.00)	3.50	17	0.003*	0.82

Note. Face donors' reported better health (i.e., less severe sickness symptoms) at the time of their healthy face photo (left) compared to the time of their sick face photo (right). For each symptom, donors reported None (0), Mild (1), Moderate (2), or Severe (3). Health rating data from 5 donors were unavailable, as reported above. Three donors did not provide health data, and one additional donor did not provide health data associated with their sick face. Means (*M*) and standard deviations (*SD*) are reported for all available symptom scores. Degrees of freedom (*df*) and Cohen's *d* effect sizes are reported for each paired samples *t* test, #*ps* < 0.05, **ps* < 0.0036 (Bonferroni correction: α_{crit} = 0.05/14).

image of that person's face when they were healthy, presented side by side, horizontally and vertically centering each face on its respective left or right side of the screen, such that the face images were spatially separated by 276–349 pixels (*M* = 317.45, *SD* = 20.48), which was 11.00–13.91 cm and 10.48–13.22 degrees. Each face was sized 291–364 pixels wide (*M* = 322.23, *SD* = 20.51) × 414–536 pixels tall (*M* = 466.82, *SD* = 35.41), which was 11.03–13.78 × 15.66–20.19 degrees (11.59–14.50 × 16.50–21.36 cm).

As in Experiment 1, we confirmed that the low-level features of our sick and healthy faces did not differ using the Saliency Toolbox (www.saliencytoolbox.net) (Ho-Phuoc et al., 2010; Walther and Koch, 2006).

3.1.3. Eye tracking system

We displayed the face photos on a 58.4 cm monitor (28 × 51 cm) with integrated eye tracking technology. We remotely tracked participants' eye gaze via corneal reflection using a Tobii TX300 eye tracker (Tobii Technology, Danderyd, Sweden) while the participant was seated 60 cm from the screen. We used a 9-point calibration and visual inspection to confirm adequate calibration.

3.1.4. Procedure

After obtaining informed consent, we calibrated participants using a 9-point calibration, instructing participants to watch the images on the screen. Then, for the experimental task, participants were instructed to watch the screen as if they were watching TV. They viewed 22 trials, each with a sick-healthy face pair displayed for 10 s, with the side of the sick and healthy faces counter-balanced to appear equally on the left and right (Fig. 3B). At the end of the study, participants reported demographic information. No participants recognized any of the people in the photos. In total, the study took approximately 5 min. Data were collected between September and October 2021.

We used Tobii Studio software (Tobii Technology, Danderyd, Sweden) to collect and summarize the data. We created oval shaped areas of interest (AOIs) around each face. AOIs were sized 360–435 pixels wide (*M* = 392.27, *SD* = 21.42; range: 14.34–17.33 cm) × 485–620 pixels tall (*M* = 392.27, *SD* = 21.42; range: 19.32–24.70 cm), slightly extended over the edges of the stimuli. We also created circle AOIs around each eye, sized 139 pixels (5.54 cm) wide × 139 pixels (5.54 cm) tall (area = 15,174.68 pixels = 24.09 cm^2), triangle AOIs around the nose, sized 190 pixels (7.57 cm) wide × 60 pixels (2.39 cm) tall (area = 15,200 pixels = 9.05 cm^2), and rectangle AOIs the mouth, sized 190 pixels (7.57 cm) wide × 80 pixels (3.19 cm) tall (area = 15,200 pixels = 24.13 cm^2), for each face (Fig. 3C). The sizes and positions of AOIs were identical within each sick-healthy face pair. AOIs around the eyes, nose, and mouth were the same size within each face pair, but varied across pairs, allowing us to use more specific shapes (circle, triangle, rectangle) while maintaining a similar area around each face region. We extracted two measures from each AOI: latency to first fixation and total fixation duration.

3.1.5. Analysis

As in Experiment 1, we first confirmed that our donors had more symptoms of illness when they contributed their sick face photo than when they contributed their healthy face photo by conducting paired samples *t* tests using two measures of overall health (Table 6).

For our primary analyses, we conducted two linear mixed effects models to compare how quickly (Prediction 2a) and how long (Prediction 2b) participants looked at the sick versus healthy faces. We also conducted a multilevel correlation to explore the association between looking times to the faces and donors' health (Prediction 2c). We then further looked at the specific face regions with linear mixed effects models to compare how quickly (latency to first fixation; Prediction 2d) and how long (total fixation duration; Prediction 2e) participants looked at the sick versus healthy eyes, noses, and mouths, with Bonferroni corrections for each measure (α_{crit} = 0.05/3 = 0.0167). We then conducted multilevel correlations to explore the associations between looking times to the three face regions—eyes, nose, and mouth—and donors' health (Prediction 2f) with Bonferroni corrections (α_{crit} = 0.05/3 = 0.0167).

All linear mixed effects models included a fixed effect for health condition and random effects of participants for health condition (Predictions 2a, 2b, 2d, and 2e). In all of the multilevel correlations (Predictions 2c and 2f), we used the difference in time looking at the sick faces, eyes, noses, and mouths relative to the healthy faces, eyes, nose and mouths (i.e., time looking at the healthy faces was subtracted from time looking at the sick faces for each AOI) and explored whether these were correlated with the differences in health reported by the face donors (i.e., scores for the healthy faces were subtracted from scores for the sick faces), for each measure of health (CCQ and SicknessQ).

3.1.6. Experiment 2 results

We detected no differences in how quickly participants looked at the sick compared to healthy faces (Table 7), suggesting no differences in overall face attention capture (Prediction 2a). However, participants looked longer to the healthy faces (*M* = 4.00 sec, *SD* = 0.38 sec) than to the sick faces (*M* = 3.83 sec, *SD* = 0.45 sec), *b* = 0.17, $t(50)$ = 2.67, *p* = 0.010, *d* = 0.72 (Fig. 3D-E; Supplementary Video 1), suggesting greater attention holding for healthy faces (Prediction 2b).

We next explored whether the *degree* of sickness was associated with differential looking to the face (Prediction 2c). We found positive correlations between the difference in health between the face images (the sicker the sick face, in terms of the number and severity of symptoms), and the looking time to the sick face (relative to the healthy face), using the CCQ, *r* = 0.10, *p* = 0.010 (Table 8, left columns; Fig. S1 Part A).

Table 7
Looking to healthy and sick faces in Experiment 2.

	Healthy Faces	Sick Faces	Fixed Effects				
	***M* (*SD*)**	***M* (*SD*)**	***b***	***t***	***df***	***p***	***d***
Latency to First Fixation							
Face	0.88 (0.48)	0.93 (0.52)	-0.004	-0.09	1320	0.932	0.01
Eyes	2.68 (0.97)	2.81 (0.99)	-0.18	−2.29	305	0.023$^{\#}$	0.27
Nose	1.85 (0.93)	1.86 (0.83)	-0.06	-0.53	50	0.598	0.15
Mouth	3.68 (1.16)	3.72 (1.23)	-0.09	-0.81	794	0.420	0.06
Total Fixation Duration							
Face	4.00 (0.38)	3.83 (0.45)	0.17	2.67	50	0.010*	0.72
Eyes	1.83 (0.67)	1.71 (0.64)	0.12	2.45	1278	0.015*	0.14
Nose	1.14 (0.59)	1.11 (0.59)	0.03	1.05	2144	0.293	0.05
Mouth	0.50 (0.32)	0.48 (0.04)	0.03	1.33	122	0.186	0.24

Note. Participants spent more time (seconds) looking at the face and eyes of healthy faces compared to sick faces, *ps < 0.0167 (Bonferroni correction: $\alpha_{crit} = 0.05/3$), and a trend of faster looking to the eyes of healthy compared to sick faces, $^{\#}$*ps* < 0.05. Means (*M*), standard deviations (*SD*), degrees of freedom (*df*), and Cohen's *d* effect sizes are reported.

Table 8
Multilevel correlations between looking time difference score and donor-reported health difference score in Experiment 2.

	CCQ		SicknessQ	
	r (*df*)	*p*	*r* (*df*)	*p*
Looking to the Face	0.10 (647)	0.010*	0.04 (647)	0.334
Looking to the Eyes	0.12 (647)	0.003*	0.004 (647)	0.909
Looking to the Nose	0.09 (647)	0.024$^{\#}$	0.06 (647)	0.115
Looking to the Mouth	-0.05 (647)	0.192	-0.01 (647)	0.842

Note. Differences in time (seconds) looking at the sick vs. healthy face, eyes, noses, and mouths, and their associations with differences in donors' reported health between sick vs. healthy, as measured by the Common Cold Questionnaire (CCQ; left) and Sickness Questionnaire (SicknessQ; right). Higher difference scores for looking indicate greater relative looking to the sick face compared to the healthy face (i.e., time looking at the healthy faces was subtracted from time looking at the sick faces). Higher CCQ and SicknessQ difference scores indicate more relative sickness at the time they donated the sick face photo relative to the healthy face photo (i.e., scores for the healthy faces were subtracted from scores from the sick faces). Positive correlations indicate that more looking at the sick faces and sick eyes were associated with sicker sick faces, *ps < 0.0167 (Bonferroni correction: $\alpha_{crit} = 0.05/3$), and a trend of more looking at the sick face noses was associated with sicker faces, $^{\#}$*ps* < 0.05.

This positive correlation indicates that the sicker sick faces were relative to other sick faces, the more looking there was at the sick faces. We did not find this relation when using the SicknessQ (Table 8, right columns), potentially suggesting our health measures may differ in their sensitivity.

Participants were faster to look at the eye region for healthy faces ($M = 2.68$, $SD = 0.97$) than sick faces ($M = 2.81$, $SD = 0.99$), $b = -0.18$, $t(305) = 2.29$, $p = 0.023$, $d = 0.27$ (Prediction 2d). However, this was not statistically significant after the Bonferroni correction. Participants looked longer to the eyes of healthy faces ($M = 1.83$ sec, $SD = 0.67$ sec) compared to the eyes of sick faces ($M = 1.71$ sec, $SD = 0.64$ sec), $b = 0.12$, $t(1278) = 2.45$, $p = 0.015$, $d = 0.14$ (Fig. 3D-E; Prediction 2e), but did not show differential looking to the nose or mouth regions (Table 7).

Lastly, we explored whether the degree of sickness was associated with differential looking to the eyes, nose, and mouth (Prediction 2f). Greater looking to the eye and nose regions of healthy faces relative to sick faces was associated with higher CCQ scores, $r(647) = 0.12$, $p = 0.003$, and $r(647) = 0.09$, $p = 0.024$, respectively (Table 8; left columns; Fig. S1 Parts B and C), but not SicknessQ scores (Table 8; right columns). However, the association between looking at the nose region and the CCQ scores did not reach statistical significance after we applied a Bonferroni correction. We detected no association between donor-reported health and attention to the mouth region (Fig. S1, Part D).

In sum, even though we found an overall effect of greater looking to the healthy faces and eyes relative to sick faces and eyes, the relative differences in health among the faces revealed an interesting pattern of greater looking associated with a sicker sick face (relative to other sick faces), underscoring the importance of considering the level of health/illness within each face.

3.1.7. Experiment 2 discussion

In a free-viewing task, when health was task-irrelevant, we found that while people did not show differences in orienting speed as a function of health (Prediction 2a), they did look longer to healthy than sick faces (Prediction 2b). We also found that differences in health *within* individuals–the same people, when acutely sick compared to when healthy–differentially influences how long people look at faces (Prediction 2c). Together, these findings suggest that people may not need to be aware of a person being potentially sick to rapidly and automatically discriminate between sick and healthy faces. People may be more willing to interact with healthy people, even when they are not prompted to consider health.

We found faster and longer looking to the healthy eye region relative to the sick eye region (Prediction 2d), but no differential looking to the nose and mouth regions (Prediction 2e). However, the faster looking to the healthy eyes was not statistically significant with the Bonferroni correction, so a replication study will be important. We also observed that the degree of facial health was positively associated with the duration of time looking to the eyes, as we predicted (Prediction 2f). Together, these findings highlight the importance of the eyes as an informative feature about health, which is consistent with the hypothesis that lassitude expressions (e.g., drooping eyelids, redder eyes) may provide nuanced information about health status (Axelsson et al., 2018; Schrock et al., 2020). These findings suggest there may be both attention capture and attention holding effects related to facial health, driven by the eye region.

Similarly, we found that the degree of facial health was positively associated with looking to the nose (Prediction 2f), though the association was not statistically significant after the Bonferroni correction. Nonetheless, these findings are in line with proposals that the nose may be an informative facial feature related to health (Manzoor and Latifi, 2021; Widen et al., 2013). Future studies are needed to systematically present key facial features, alone or in combination, to further uncover which parts of the face are used—individually or in combination—to identify health.

Our finding of preferential looking to healthy faces is in line with reports of longer looking to more attractive faces (Glaholt et al., 2009; Hönekopp, 2006) and faces with whom people are drawn to interact (Aharon et al., 2001; Leder et al., 2010; Sui and Liu, 2009). Indeed, attractive individuals are perceived as healthier (Kalick et al., 1998) and are more likely to elicit social approach (Lemay et al., 2010; Waynforth, 2001). Our findings that degree of health is positively associated with looking patterns suggests that humans may be drawn to look longer at healthier people, which may reflect a sensitivity to health that facilitates evaluation of potentially risky social interactions (Gul et al., 2021; Nunner et al., 2021). Further, given that people discriminated healthy from sick faces without external features (e.g., hair), our findings demonstrate that the inner facial features, in particular, provide information about health. This result suggests that our findings in Experiment 1 are also likely driven by aspects of the faces themselves, rather than differences in the outer elements, such as hairstyles.

While longer looking to healthy relative to sick faces may reflect a draw to socially approach, an alternative interpretation is that people may be avoiding sick faces. Lassitude expressions, which convey illness, may communicate to others an unwillingness to engage in social interactions (Schrock et al., 2020). As such, people may avoid such individuals, because they do not appear to be friendly or open to having social interactions. Indeed, when viewing faces, people report being less inclined to socialize with healthy people who have been sleep-deprived (Sundelin et al., 2017) or who have negative expressions (Nikitin and Freund; 2019), both of which are qualities that also vary with health, as we found in Experiment 1. There is, therefore, a need to better understand the mechanisms involved in recognizing and avoiding unhealthy conspecifics.

4. Experiment 3

One approach to begin to understand the mechanisms involved in sick face perception is to track participants' physiological arousal. Pupil dilation is an unobtrusive measure of physiological arousal (Bradley et al., 2008; Wang et al., 2018) that can test competing hypotheses. If our findings in Experiment 2 were driven by an avoidance of sick faces, then we expect greater arousal (more pupil dilatation) when observing sick faces—reflecting threat detection—compared to when observing healthy faces (Prediction 3a; Kret et al., 2013). In contrast, if our findings reflect an increased desire to approach healthy faces, we expect greater arousal (more pupil dilation) when observing healthy faces (Prediction 3b)—reflecting an appetitive reaction (Kuraguchi and Kanari, 2021)—compared to when observing sick faces.

Another aim of Experiment 3 was to explore how participants' approach-avoidance responses occur in simulated "real world" social encounters. We described a common scenario in which there is a high likelihood of encountering people who are sick: entering a crowded waiting room at a doctor's office. We did this to increase disease salience and perceived infectability, which may increase preferences for health (Brown and Sacco, 2022; Brown et al., 2021). We predicted that participants' pupil dilation would be associated with their approach ratings in one of two ways: either desire to approach would be negatively associated with arousal, reflecting threat detection and avoidance (Prediction 3c), or desire to approach would be positively associated with arousal, reflecting an appetitive reaction (Prediction 3d). We expected to replicate our Experiment 1 findings that participants' approach ratings would be negatively associated with facial health, i.e., less approach to sicker faces (Prediction 3e).

Finally, the speed with which approach-avoidance decisions are made can inform whether such decisions are driven more by appetitive or avoidance motivations (Smillie and Jackson, 2005). Because it is generally safer to avoid unfamiliar people than it is to approach them (Kurzban and Leary, 2001; Schaller et al., 2003, 2015), avoidance decisions may occur more rapidly than approach decisions (Chen and Bargh, 1999). We, therefore, predicted more rapid decisions for sicker faces (Prediction 3f). We also predicted slower decision speeds would be associated with higher approach ratings (reflecting increased caution and hesitation to approach; Predictions 3g). Finally, if greater pupil dilation reflects avoidance of sick faces, we predicted faster decision speeds would be associated with greater pupil dilation (Prediction 3h; indicating more rapid decisions to avoid).

4.1. Materials and methods

4.1.1. Participants

We recruited new participants ($N = 112$), who did not participate in Experiments 1 or 2. Participants received course credit for their participation. The University of Miami Institutional Review Board approved this study.

4.1.2. Stimuli

We used 64 neutral/relaxed face photos collected from 32 individuals, once when sick and once when healthy (Fig. 3A). These included the original 44 images from 22 individuals from Experiment 2, as well as 20 additional images from 10 new individuals (see Table S1). Among the 10 new individuals 1 was wearing glasses in both their healthy and sick photos. As in Experiments 1 and 2, we obtained consent to use face photos and collected health information from face donors (Table 9). Sick and healthy photos were taken an average of 4.81 months apart. We confirmed that donors had more severe symptoms (e.g., fever) when sick, reflected in a higher SicknessQ (Andreasson et al., 2018) total score when they were sick ($M = 16.78$, $SD = 6.57$) compared to when they were healthy ($M = 1.59$, $SD = 2.29$), $t(26) = 12.11$, $p < 0.001$, $d = 2.33$, and a higher CCQ (Powell et al., 2008) total score when they were sick ($M = 15.93$, $SD = 5.89$) compared to when they were healthy ($M = 1.48$, $SD = 2.67$), $t(26) = 12.57$, $p < 0.001$, $d = 2.42$. CCQ and SicknessQ scores positively correlated with one another, $r(54) = 0.91$, $p < 0.001$, confirming that they were both capturing the degree of illness/health.

As in Experiment 2, in Experiment 3 we cropped each image in an oval shape (see Fig. 3A for examples). We displayed faces one at a time with the same person's face shown twice, once when sick and once when healthy. Each face was sized 300–399 pixels wide ($M = 337.06$, $SD = 22.88$) × 428–641 pixels tall ($M = 499.50$, $SD = 46.40$), which was 11.37–31.10 × 16.17–24.03 degrees (11.95–33.39 cm × 17.05–25.54 cm).

Table 9
Experiment 3 donor-reported symptoms and severity for healthy and sick face photos.

	Healthy Faces	Sick Faces				
	***M* (*SD*)**	***M* (*SD*)**	*t*	*df*	*p*	*d*
Chills	0.00 (0.00)	1.00 (0.98)	5.50	26	< 0.001*	1.06
Headache	0.21 (0.63)	1.79 (1.17)	6.81	26	< 0.001*	1.31
Joint pains	0.14 (0.45)	0.96 (1.09)	4.23	25	< 0.001*	0.84
Muscle or body aches	0.07 (0.26)	1.50 (1.07)	7.32	26	< 0.001*	1.41
Fatigue or tiredness	0.39 (0.57)	2.46 (0.84)	11.76	26	< 0.001*	2.26
Nausea	0.07 (0.38)	0.93 (1.05)	4.40	26	< 0.001*	0.85
Vomiting	0.00 (0.00)	0.18 (0.61)	1.55	26	0.134	0.30
Diarrhea	0.00 (0.00)	0.36 (0.87)	1.98	26	0.059	0.38
Abdominal pain	0.00 (0.00)	0.57 (0.879)	3.47	26	0.002*	0.67
Sneezing	0.11 (0.32)	1.04 (1.07)	4.35	26	< 0.001*	0.84
Sore throat	0.00 (0.00)	1.64 (1.13)	7.37	26	< 0.001*	1.42
Nasal discharge	0.21 (0.50)	1.00 (1.02)	3.85	26	< 0.001*	0.74
Nasal obstruction	0.11 (0.32)	0.89 (1.10)	4.08	26	< 0.001*	0.78
Cough	0.11 (0.32)	1.07 (1.02)	4.35	26	< 0.001*	0.84

Note. Face donors' reported significantly better health (i.e., less severe sickness symptoms) at the time of donating their healthy face photo (left) compared to the time they donated their sick face photo (right). For each symptom, donors reported None (0), Mild (1), Moderate (2), or Severe (3). Health rating data from 5 donors were unavailable, as reported above. Three donors did not provide health data, one additional donor did not provide health data associated with their sick face, and one additional donor did not provide health data associated with their healthy face. Means (*M*) and standard deviations (*SD*) are reported for all available symptom scores. Degrees of freedom (*df*) and Cohen's d effect sizes are reported for each paired samples *t* test, *$ps < 0.0036$ (Bonferroni correction: $\alpha_{crit} = 0.05/14$).

As in Experiment 1, we confirmed that low-level features of our sick and healthy faces did not differ using the Saliency Toolbox (see Supplementary Method; https://www.saliencytoolbox.net; Ho-Phuoc et al., 2010; Walther and Koch, 2006). In addition, to increase experimental control in our stimuli, we used the SHINE_color toolbox (Dal Ben, 2021a; 2021b; Willenbockel et al., 2010) to match images on their luminance.

4.1.3. Eye tracking system

As in Experiment 2, we displayed face photos on a 58.4 cm monitor and tracked pupil dilation using a Tobii TX300 eye tracker (Tobii Technology, Danderyd, Sweden).

4.1.4. Procedure

We explored participants' physiological responses and approach-avoidance ratings in a simulated "real world" social encounter. After obtaining informed consent, we calibrated participants using a 9-point calibration, instructing participants to watch the images on the screen while imagining searching for a seat in a crowded waiting room of a doctor's office. They were told, "Imagine you are going to the doctor. When you arrive, you enter a crowded waiting room and you're looking for a seat. There are some seats available, but because the room is full, you have to be near someone. We're going to show you a series of faces of these people and you have to decide how likely you'd be to sit next to each one." Then, participants viewed a single face image for 5 s (Fig. 4B), followed by a question about the person that appeared on the screen: "How likely would you be to sit next to this person…." They responded using a keyboard, on a scale from 1 (very unlikely) to 9 (very likely). They viewed 64 randomized trials. Each trial began with a 1-second display of a black fixation cross on a white background. At the end of the study, participants reported demographic information. No participants recognized any of the people in the photos. In total, the study took approximately 10 min. Data were collected between April and December 2022.

4.1.5. Pupil dilation measure

Raw pupil diameters (sampling rate: 300 HZ) were extracted from Tobii Studio. All pupil data preprocessing was conducted in RStudio. First, for data that were missing from one eye, we interpolated missing samples from the other eye, given the high correlation between left and right pupil sizes (Jackson and Sirois, 2009), and calculated the average pupil diameters from both eyes. Unreliable pupil samples were removed. Removed samples included outliers of pupil diameters ($<$1.5 mm or $>$ 8 mm), speed outliers (speed of diameter change between consecutive samples detected using median absolute deviation), and outliers from trend lines (trend lines of general additive models of time; outliers detected using median absolute deviation). We performed a linear interpolation to fill in gaps of missing data that were smaller than 100 ms (Geangu et al., 2011; Jackson and Sirois, 2009; Kret and Sjak-Shie, 2018). The preprocessed pupil data were down-sampled to 20 Hz (aggregated to 50 ms time-bins) and smoothed with the average of rolling windows of 5 time-bins (van Reekum et al., 2007).

Individual baseline pupil sizes were defined as the average pupil sizes in the fixation cross period preceding each face stimulus onset. Pupil dilation for each time-bin while looking at the faces was calculated as the difference in pupil diameters from the baseline pupil size to correct for individual differences in pupil sizes (Geangu et al., 2011; Gredebäck and Melinder, 2010; Hepach and Westermann, 2016).

4.1.6. Behavioral measures

Participants rated each face on a scale from unlikely to approach (1) to very likely to approach (9). We recorded their key press latencies as the duration of time between when the rating scale appeared on the screen and the key press, reflecting the speed with which participants made their approach-avoidance decisions.

4.1.7. Analysis

We conducted five bootstrapped cluster-based permutation analyses (1000 bootstrap iterations) with linear mixed effects models to examine whether the time-series change in pupil dilation over the time course of viewing the face images varied as a function of facial health (face type: sick or healthy), donor-reported levels of health (CCQ and SicknessQ scores), participants' approach ratings, and participants' key press latencies (Predictions 3a-3d). Each linear mixed effects model included time-series pupil dilation as the outcome variable and each of the health conditions of the faces, donor-reported health, the participants' approach ratings, and key press latencies, respectively, as the predictors. Our limited sample size of face donors (N = 32) might lead to a lack of power and an increased estimation bias for the donor-level random effects (Maas and Hox, 2005; Westfall et al., 2014); therefore, we only controlled for random effects of participants for each predictor in the models. The bootstrapped cluster-based permutation analysis requires statistical significance to be achieved in two parts: the first part reveals statistically significant time cluster(s) using linear mixed effects modeling at each timepoint, followed by a second part that consists of a bootstrapped permutation test for these clusters to control for inflated Type-1 errors due to analyses at multiple timepoints. Compared to traditional functional data analyses, the bootstrapped cluster-based permutation analysis is compatible with any null hypothesis testing statistics on time-series data and controls the false-alarm rate without sacrificing much statistical power (Franzen et al., 2022; Maris and Oostenveld, 2007).

We explored whether participants' approach ratings varied as a function of donor-reported health (Prediction 3e). First, we conducted a linear mixed effects model with a fixed effect for health condition and random effects of participants for health condition comparing sick and healthy faces. We also conducted two multilevel correlations controlling for random effects of participants to explore the association between participants' approach ratings and donors' health scores using the CCQ and the SicknessQ.

Finally, we conducted four multilevel correlations controlling for random effects of participants for key press latencies to explore participants' response speeds in relation to face donors' health (CCQ and SicknessQ scores; Prediction 3f), approach ratings (Prediction 3g), and pupil dilation (Prediction 3h). We applied Bonferroni corrections for tests that included multiple measures (see Tables 10 and 11 for details).

4.1.8. Experiment 3 results

The permutation analysis revealed a statistically significant time cluster from 2.40 to 3.10 s post stimulus onset that showed greater pupil dilation to the sick faces compared to the healthy faces (summed t statistics = 38.66, p = 0.033; Fig. 4D-F), in line with Prediction 3a (and inconsistent with Prediction 3b). This finding suggests that participants may have differential physiological reactions to faces of varying health. However, pupil dilation was not associated with either the SicknessQ or the CCQ at any time point during face viewing (see Figures S2 and S3 for details), which may indicate that, in this task, we may have lacked sensitivity to capture physiological responses to subtly varying degrees of sickness.

However, we did find that pupil dilation was negatively associated with the participants' approach rating from 0.6 to 2.65 s (summed t statistics = −111.60, p < 0.001) and from 2.80 to 4.95 s (summed t statistics = −124.92, p < 0.001), which was 88% of the time that the stimulus was on the screen (Fig. 5), consistent with Prediction 3c (and failing to support Prediction 3d). This finding supports our hypothesis that physiological arousal may be associated with avoidance of potential contagion threats.

Participants reported that they were more likely to approach (i.e., sit next to) people with healthy faces (M = 5.46, SD = 2.20) compared to sick faces (M = 5.28, SD = 2.21), b = 0.18, SE = 0.05, t(4000) = 3.85, p < 0.001, d = 0.12 (Fig. 4C; Table 10; Prediction 3e).

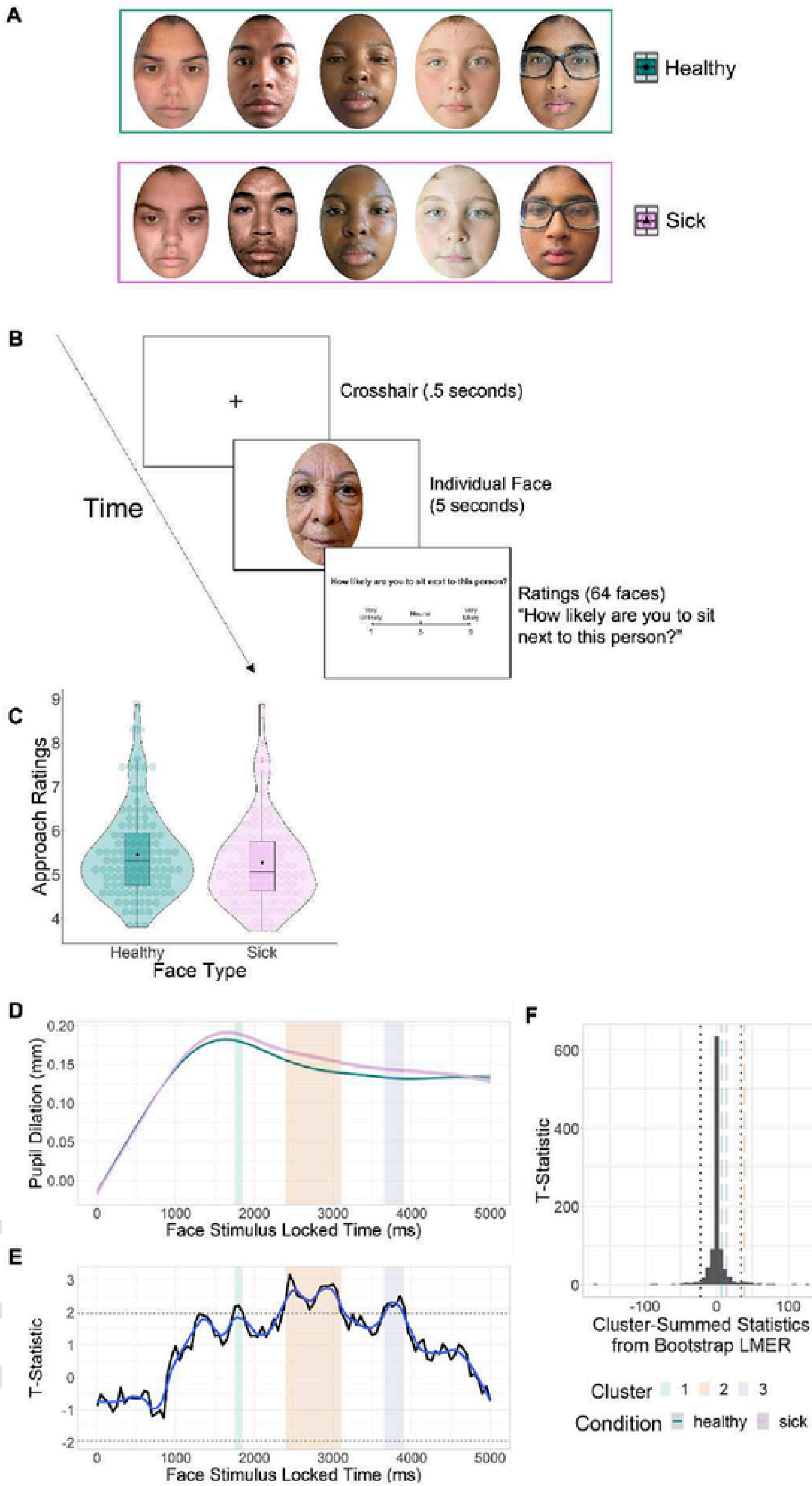

Fig. 4. (A) Healthy (top row; teal blue) and sick (bottom row; pink) oval-cropped face stimuli examples. (B) Example of study procedure. (C) Approach ratings to the healthy faces (teal blue, circles) and sick faces (pink, triangles). Horizontal lines within the boxplots indicate the medians. Black circles/triangles indicate means. Hinges of the boxplots show the first (bottom) and third (top) quartiles. The whiskers extend up to 1.5 × Interquartile Range, above and below the hinges. Dots are binned raw data (e.g., each individual participant's rating). Violin plots show the frequency distributions of the ratings and peak pupil dilations. (D) Pupil

dilation to healthy (teal blue curve) and sick (pink curve) faces over time in milliseconds (ms; gray ribbon = standard errors). Time clusters highlighted in color indicate statistically significant divergence. (E) The t statistics for the differences in pupil dilation to sick versus healthy faces (black solid line = raw statistics; blue solid line = LOESS smoothed statistics) vary as a function of time (x-axis). The dashed horizontal lines indicate the two-sided critical *t* values at $\alpha = 0.05$ for linear mixed effects model ($df = 3233$). Time clusters highlighted in color indicate statistically significant divergence at the time point level with the linear mixed effects model. (F) Distribution of summed statistics of the permutation test with 1000 bootstrap iterations. Vertical black dotted lines = 95% chance levels. Vertical dashed colored lines = summed t statistics for the three time clusters. Only Cluster 2 was beyond chance levels (statistically significant) after the permutation test (Cluster 1: $p = 0.257$; Cluster 2: $p = 0.033$; Cluster 3: $p = 0.120$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 10
Experiment 3 Healthy and Sick Face Differences Across Measures.

	Healthy		Sick						
	M	*SD*	*M*	*SD*	*b*	*t*	*df*	*p*	*d*
Approach Rating	5.45	2.20	5.28	2.21	0.17	3.68	4000	< 0.001*	0.12
Key Press Latency	1.40	0.99	1.36	0.99	0.03	1.73	4527	0.083	0.05
Peak Pupil Dilation	0.39	0.25	0.39	0.25	-0.004	0.75	3666	0.455	0.02
SicknessQa	1.59	2.29	16.78	6.57	–	12.11	26	< 0.001*	2.33
CCQa	1.48	2.67	15.93	5.89	–	12.57	26	< 0.001*	2.42

Note. Means (*M*), standard deviations (*SD*), degrees of freedom (*df*), and Cohen's d effect sizes calculated from the linear mixed effects models are reported for each outcome variable comparing healthy faces (left) and sick faces (right). Approach ratings reflect participants' reports of how likely they were to sit next to someone in a doctor's office. Key press latencies reflect how quickly they made these ratings. Peak pupil dilation reflects their peak autonomic arousal during the 5 s of viewing each face. Common Cold Questionnaire (CCQ) and Sickness Questionnaire (SicknessQ) scores captured donor-reported health, with greater scores indicating more severe symptoms. aCCQ and SicknessQ were compared using paired samples t tests given that there was no participant-level data for them. *$ps < 0.01$ (Bonferroni correction: $\alpha_{crit} = 0.05/5$).

Furthermore, we found that the degree of health—measured by the CCQ and SicknessQ (number/severity of symptoms)—were also negatively correlated with their approach ratings, $r(6236) = -0.06$, $p < 0.001$, and $r(6236) = -0.03$, $p = 0.025$, respectively (Table 11). That is, participants were more likely to approach healthier faces. However, the correlation between CCQ and the approach ratings was not statistically significant after Bonferroni correction. Together, these results suggest that people can distinguish sick from healthy faces, overall, but may not have a very accurate perception of the degree of health of any given face.

We did not observe differences in decision speed (i.e., key press latency) between healthy ($M = 1.40$, $SD = 0.99$ s) and sick faces ($M = 1.36$, $SD = 0.99$ s), $b = 0.03$, $SE = 0.02$, $t(4527) = 1.73$, $p = 0.083$, $d = 0.05$. However, we found a negative correlation between decision speed and degree of sickness (number/severity of symptoms): CCQ scores, $r(6236) = -0.04$, $p = 0.001$, SicknessQ scores, $r(6236) = -0.03$, $p = 0.008$. Together, these findings are consistent with our prediction that participants more rapidly rate sicker faces (Prediction 3f).

However, we detected no correlation between decision speed and approach ratings, $r(7010) = 0.01$, $p = 0.306$, failing to support our prediction that there may be more caution and hesitation to approach someone than to avoid someone (Prediction 3g). We also detected no association between pupil dilation and decision speed at any time point (see Supplementary Fig. S4; Prediction 3h).

4.1.9. Experiment 3 discussion

We found elevated physiological arousal—reflected in greater pupil dilation—to sick faces compared to healthy faces, consistent with a threat avoidance response (Prediction 3a). Indeed, we also found greater pupil dilation was associated with lower approach ratings (higher avoidance ratings), in line with Prediction 3c and similar to our results in Experiment 1 in which we found that lower approach ratings were associated with greater donor-reported sickness (Prediction 1b) and greater sickness ratings (Prediction 1c). Together, our findings are in line with the hypothesis that sick faces may activate heightened arousal and avoidance. Further support for this hypothesis comes from a report that systolic blood pressure increased—reflecting increased arousal—in response to salient tactile and auditory cues of disgust (e.g., sound of coughing; Croy et al., 2013).

In addition to analyzing pupil dilation data with a time-series analysis, we also explored how peak pupil dilation—an indicator of the maximum level of physiological arousal—was associated with the health of a face (see Supplementary Materials). Consistent with our time-series analysis, neither CCQ nor SicknessQ scores reflecting facial health were correlated with peak pupil dilation.

We also replicated our findings of approach-avoidance ratings (from Experiment 1) using a more realistic simulated "real world" social encounter scenario–deciding who to sit next to at the doctor's office. We found, in this scenario, people were sensitive to facial health, choosing to sit next to people who were healthier and avoiding sitting next to those who were sicker (Prediction 3e). These findings are similar to a study in which 8- to- 9-year-old children reported that they were more likely to sit next to a person with a healthy face compared to a sick face (Leung et al., under review). Together, these studies suggest that humans may be sensitive to facial cues of disease/health already in childhood and continuing into adulthood, which may promote behavioral

Table 11
Multilevel correlations controlling for random effects of participants among experiment 3 behavioral and physiological measures.

	Key Press Latency	Peak Pupil Dilation	SicknessQ	CCQ
Approach Rating	$r(7010) = 0.01$, $p = 0.306$	$r(6004) = -0.05$, $p < 0.001$*	$r(6236) = -0.06$, $p < 0.001$*	$r(6236) = -0.03$, $p = 0.025$#
Key Press Latency		$r(6004) = -0.03$, $p = 0.008$#	$r(6236) = -0.03$, $p = 0.008$#	$r(6236) = -0.04$, $p < 0.001$*
Peak Pupil Dilation			$r(5353) = -0.02$, $p = 0.230$	$r(5353) = -0.04$, $p = 0.012$#
SicknessQ				$r(54) = 0.91$ $p < 0.001$*

Note. Multilevel correlations among the four dependent measures. Approach ratings reflect participants' reports of how likely they were to sit next to someone in a doctor's office. Key press latencies reflect how quickly they made these ratings. Peak pupil dilation reflects their peak autonomic arousal during the 5 s of viewing each face. Common Cold Questionnaire (CCQ) and Sickness Questionnaire (SicknessQ) captured donor-reported health, with greater scores indicating more severe symptoms. The association between the CCQ and SicknessQ was conducted using a Pearson correlation given that there were no participant-level data for them (health data were at the face stimulus level only). *$ps < 0.005$ (Bonferroni correction: $\alpha_{crit} = 0.05/10$), #$ps < 0.05$.

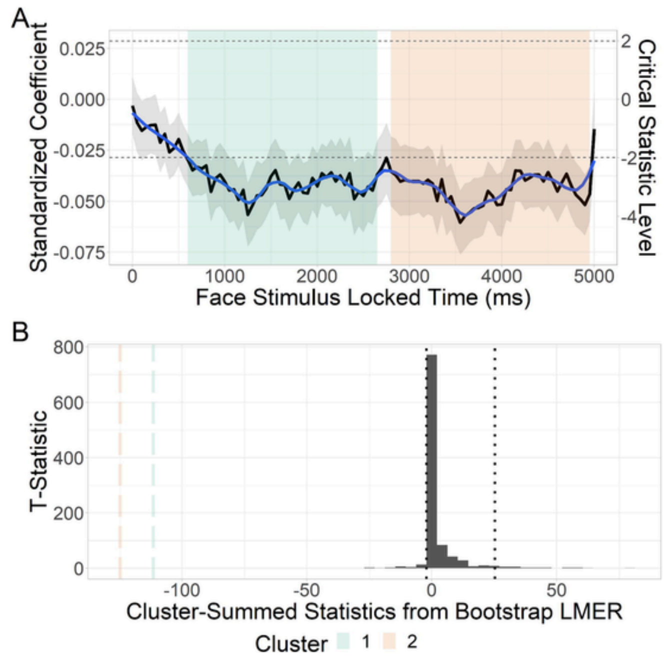

Fig. 5. (A) The standardized coefficients between pupil dilation to the face stimuli (y-axis) and participants' approach ratings (black solid line = raw coefficients; blue solid line = LOESS smoothed coefficients; gray ribbon = standard errors) varied as a function of time (x-axis) in milliseconds (ms). The dashed horizontal lines indicate the two-sided critical *t* values at $\alpha = 0.05$ for linear mixed effects model ($df = 71.50$). Time clusters highlighted in colors (green and orange) indicate statistically significant associations between pupil dilation and approach ratings given the time point-level linear mixed effects model. (B) Distribution of summed statistics of the permutation test with 1000 bootstrap iterations. Black dotted lines = 95% chance levels. Long dashed lines = summed t statistics for clusters in (A). Clusters beyond chance levels are statistically significant after the permutation test ($ps < 0.001$ for both clusters). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

avoidance of potential threats (Hugenberg and Bodenhausen, 2003; Roesmann et al., 2020).

Participants' faster decision speeds (how quickly they completed the ratings) were associated with sicker faces (Prediction 3f). These findings suggest that the speed with which participants were making their decisions was influenced by the health of the people who they were deciding to approach or avoid. Participants were free to respond as quickly or slowly as they wanted, so it is surprising that we observed differences in response speed related to health. This result may suggest that decisions related to approaching strangers, rather than avoiding them, require more careful cost-benefit analysis (Kurzban and Leary, 2001; Schaller et al., 2003, 2015).

5. General discussion

Pathogens are a leading cause of death world-wide (World Health Organization, 2020). Given this extraordinary selective pressure, humans are theorized to have evolved abilities to detect and avoid contagious diseases (e.g., bodily fluids, biological motion; Curtis et al., 2004; Regenbogen et al., 2017). In three experiments, we tested whether people detect subtle cues of acute, contagious illness in the faces of people who are ill compared to the same individuals when healthy. Compared to healthy faces, sick faces were rated as sicker, more dangerous, more likely to be avoided, eliciting more unpleasant feelings, appearing more tired, and having more negative expressions (Experiment 1). Participants looked less (Experiment 2) and exhibited greater peak pupil dilation (Experiment 3) when viewing the sick faces compared to healthy faces, perhaps suggesting they may be less willing to engage with sicker people. Furthermore, across all experiments, participants' responses varied as a function of the health status reported by the face donors, suggesting a nuanced sensitivity to the *degree* of sickness. Together, these findings suggest that humans may detect subtle threats of contagion from sick faces, which may facilitate illness avoidance.

When sick, humans, like other animals, show reductions in their sociability (Devlin et al., 2022). We found people rated sick faces as being more tired and displaying more negative affect (Experiment 1); therefore, viewers may have detected that sick people had decreased desires to socialize. As a result, people may have reported a reduced interest in approaching, not necessarily to avoid illness, but to avoid a person who appears antisocial. People report less interest in socializing with experimentally sleep-deprived people, based on viewing photos of their faces (Sundelin et al., 2017), potentially suggesting that facial expressions (e.g., tiredness, negative affect), regardless of whether they are related to infectious illness, may broadly influence approach decisions. In addition, while we only measured responses to strangers (unfamiliar people), responses to familiar sick individuals, especially those with whom one is close (e.g., child, partner), may elicit more affiliative/approach responses to provide care and comfort rather than avoidance (Muscatell and Inagaki, 2021). When people are induced to feel sick (without actually being contagious), they report a greater desire to affiliate with close social partners, such as their friends and parents (Inagaki et al., 2015). Together, these studies suggest that lassitude expressions may trigger different responses in viewers as a function of their relationship.

In contrast to our prediction, we did not find that people were faster to orient to sick faces, or faster to show pupil dilation to sick faces, compared to healthy faces, as they are with other types of threats (De Oca and Black, 2013). That is, we found little evidence that there is a rapid threat detection mechanism for sick faces operating in the initial phases of attention capture. People may need additional time to extract/interpret health-relevant information in faces. This delayed response may allow for people to make more accurate decisions about whether to engage with others, thus reducing approach to dangerous pathogens and avoidance of beneficial social interactions with healthy conspecifics (Schaller et al., 2015). Alternatively, other types of tasks (e.g., visual search, gap-overlap) may better capture subtle differences in attentional processing (Kikuchi et al., 2011; Langton et al., 2008).

5.1. Limitations and future directions

The current studies relied on photo stimuli, which enabled a high degree of experimental control, but were less ecologically valid compared to live interactions. In the real world, people commonly have access to richer, dynamic, and multimodal information (e.g., body movements, voice, clothing) during social encounters, which, combined with faces, may provide more accurate insights into a person's health. Without this information, it is quite impressive that people were still sensitive to subtle cues of illness using *only static images* of faces, consistent with proposals of a robust, flexible, health-detection system (i.e., behavioral immune system). In fact, observing real-world stimuli likely leads to even more accurate sickness detection than what we observed in the current studies.

Unfortunately, we could not explore individual differences due to our modest sample sizes. Future studies can explore whether there are individual differences in illness detection, particularly among disease-vulnerable populations. For example, the behavioral immune system may be adaptively elevated among immuno-suppressed people (Fleischman and Fessler, 2011; Makhanova and Shepherd, 2020), such as pregnant people (Jones et al., 2005).

Another limitation is that the current study included only one sick face from each person. Future studies could include *multiple* sick faces, at different time points, from the same face donor, which may capture varying symptom severities. Tracking sick face perception in relation to sick people's physiological markers of infectiousness, such as viral load, could guide efforts to identify methods to improve sickness detection. Further, studies should continue to collect sick and healthy face photos from a diverse group to expand the number of stimuli. Nonetheless, the current study reflects a first step in establishing a new, sensitive method for measuring facial sickness perception.

A notable strength of the current study is that our face donors were of diverse ages, genders, races, and ethnicities. However, our sample mostly consisted of White, Non-Hispanic U.S. participants. While some aspects of (experimentally-induced) sickness in faces may be recognized across cultures (Arshamian et al., 2021), it is unclear whether our findings would generalize to other populations (e.g., outside of the U.S.) and, therefore, replication studies in other cultures will be important.

5.2. Conclusions

We found that humans detect legitimate sickness using only static facial cues. In addition to rating sick faces as being sicker and less approachable than healthy faces, people also looked less to, and demonstrated greater pupil dilation to, sick faces compared to healthy faces. Together, these findings offer additional support for the hypothesis that humans evolved an illness-detection adaptation to preemptively avoid contagion (Curtis, 2014; Regenbogen et al., 2017; Schaller and Park, 2011). Paradoxically, our human desire to affiliate during times of crisis and threat, such as during pandemics, may drive people to seek and provide comfort to others, elevating the transmission of contagious illness (Dezecache et al., 2020). Accurately detecting and avoiding illness would enable more selective affiliation, allowing people to safely obtain and provide social support during crises. A more thorough understanding of humans' pathogen avoidance tendencies is a necessary first step to potentially improve these skills, thereby slowing the spread of illness and, ultimately, improving public health (Ackerman et al., 2021).

Declaration of Competing Interest

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

Data availability

Data files are available in Supplementary Materials.

Acknowledgements

We are grateful to the people who helped us collect and prepare stimuli, including Shantalle Martinez, Isabella Cabrera, Amy Ahn, and Baylee Brochu. Thanks to Debra Lieberman and Laurie Bayet for contributions to study design. Thanks to Roberto Lazo for programming the tasks and to Hayley Nathaniel and Shirsika Kummeta for support in collecting data. Thanks to the undergraduate research assistants in the Social Cognition Laboratory for feedback on an earlier draft. Funded by National Science Foundation CAREER Award 1653737 (to EAS), Association for Psychological Science James McKeen Cattell Fund Fellowship Sabbatical Award (to EAS), and Alvin V., Jr. and Nancy C. Baird Professorship (to KVJ).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbi.2023.03.007.

References

Ackerman, J.M., Merrell, W.N., Choi, S., 2020. What people believe about detecting infectious disease using the senses. Curr. Res. Ecol. Soc. Psychol. 1, 100002. https://doi.org/10.1016/j.cresp.2020.100002.

Ackerman, J.M., Tybur, J.M., Blackwell, A.D., 2021. What role does pathogen-avoidance psychology play in pandemics? Trends Cogn. Sci. 25 (3), 177–186. https://doi.org/10.1016/j.tics.2020.11.008.

Aharon, I., Etcoff, N., Ariely, D., Chabris, C.F., O'connor, E., Breiter, H.C., 2001. Beautiful faces have variable reward value: fMRI and behavioral evidence. Neuron 32 (3), 537–551. https://doi.org/10.1016/S0896-6273(01)00491-3.

Andreasson, A., Wicksell, R.K., Lodin, K., Karshikoff, B., Axelsson, J., Lekander, M., 2018. A global measure of sickness behaviour: Development of the Sickness Questionnaire. J. Health Psychol. 23 (11), 1452–1463. https://doi.org/10.1177/1359105316659917.

Arshamian, A., Sundelin, T., Wnuk, E., O'Meara, C., Burenhult, N., Rodriguez, G.G., Lekander, M., Olsson, M.J., Lasselin, J., Axelsson, J., Majid, A., 2021. Human sickness detection is not dependent on cultural experience. Proc. R. Soc. B 288 (1954), 20210922.

Axelsson, J., Sundelin, T., Olsson, M.J., Sorjonen, K., Axelsson, C., Lasselin, J., Lekander, M., 2018. Identification of acutely sick people and facial cues of sickness. Proc. R. Soc. B Biol. Sci. 285 (1870), 20172430. https://doi.org/10.1098/rspb.2017.2430.

Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw., 67(1). 10.18637/jss.v067.i01.

Berdica, E., Gerdes, A., Bublatzky, F., White, A. J., Alpers, G. W. 2018. Threat vs. threat: Attention to fear-related animals and threatening faces. Front. Psychol., 9, 1154. 10.3389/fpsyg.2018.01154.

Bradley, M.M., Miccoli, L., Escrig, M.A., Lang, P.J., 2008. The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology 45 (4), 602–607. https://doi.org/10.1111/j.1469-8986.2008.00654.x.

Bressan, P., 2021. Strangers look sicker (with implications in times of COVID-19). Bioessays 43 (3), 2000158. https://doi.org/10.1002/bies.202000158.

Brown, M., Sacco, D.F., 2022. How and when crowd salience activates pathogen-avoidant motives. Evol. Behav. Sci. 16 (1), 23–37. https://doi.org/10.1037/ebs0000191.

Brown, M., Tracy, R.E., Young, S.G., Sacco, D.F., 2021. Crowd salience heightens tolerance to healthy facial features. Adapt. Hum. Behav. Physiol. 7 (4), 432–446. https://doi.org/10.1007/s40750-021-00176-2.

Campbell, N.M., Dawel, A., Edwards, M., Goodhew, S.C., 2021. Does motivational intensity exist distinct from valence and arousal? Emotion. https://doi.org/10.1037/emo0000883.

Centers for Disease Control and Prevention. (2021, July 1). V-safe After Vaccination Health Checker. https://www.cdc.gov/coronavirus/2019-ncov/vaccines/safety/vsafe.html.

Chen, M., Bargh, J.A., 1999. Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. Pers. Soc. Psychol. Bull. 25 (2), 215–224. https://doi.org/10.1177/0146167299025002007.

Crouzet, S.M., Kirchner, H., Thorpe, S.J., 2010. Fast saccades toward faces: face detection in just 100 ms. J. Vis. 10 (4), 16. https://doi.org/10.1167/10.4.16.

Croy, I., Laqua, K., Süß, F., Joraschky, P., Ziemssen, T., Hummel, T., 2013. The sensory channel of presentation alters subjective ratings and autonomic responses toward disgusting stimuli—Blood pressure, heart rate and skin conductance in response to visual, auditory, haptic and olfactory presented disgusting stimuli. Front. Hum. Neurosci. 7, 510. https://doi.org/10.3389/fnhum.2013.00510.

Curtis, V.A., 2014. Infection-avoidance behaviour in humans and other animals. Trends Immunol. 35 (10), 457–464. https://doi.org/10.1016/j.it.2014.08.006.

Curtis, V., Aunger, R., Rabie, T., 2004. Evidence that disgust evolved to protect from risk of disease. Proc. R. Soc. Lond. B 271 (suppl_4).

Dal Ben, R. 2021a. SHINE_color: controlling low-level properties of colorful images. 10.31234/osf.io/fec6x.

Dal Ben, R. 2021b. SHINE_color and Lum_fun: A set of tools to control luminance of colorful images (Version 0.3). [Computer program]. 10.17605/OSF.IO/AUZJY, retrieved from https://osf.io/auzjy/.

De Oca, B.M., Black, A.A., 2013. Bullets versus burgers: is it threat or relevance that captures attention? Am. J. Psychol. 126 (3), 287–300. https://doi.org/10.5406/amerjpsyc.126.3.0287.

Devlin, B.A., Smith, C.J., Bilbo, S.D., 2022. Sickness and the social brain: How the immune system regulates behavior across species. Brain Behav. Evol. 97 (3–4), 197–210. https://doi.org/10.1159/000521476.

Dezecache, G., Frith, C.D., Deroy, O., 2020. Pandemics and the great evolutionary mismatch. Curr. Biol. 30 (10), R417–R419. https://doi.org/10.1016/j.cub.2020.04.010.

Erdfelder, E., Faul, F., Buchner, A., 1996. GPOWER: A general power analysis program. Behav. Res. Methods Instrum. Comput. 28 (1), 1–11. https://doi.org/10.3758/BF03203630.

Federico, G., Ferrante, D., Marcatto, F., Brandimonte, M.A., 2021. How the fear of COVID-19 changed the way we look at human faces. PeerJ 9, e11380.

Feldmann-Wüstefeld, T., Schmidt-Daffy, M., Schubö, A., 2011. Neural evidence for the threat detection advantage: differential attention allocation to angry and happy faces. Psychophysiology 48 (5), 697–707. https://doi.org/10.1111/j.1469-8986.2010.01130.x.

Fleischman, D.S., Fessler, D.M., 2011. Progesterone's effects on the psychology of disease avoidance: support for the compensatory behavioral prophylaxis hypothesis. Horm. Behav. 59 (2), 271–275. https://doi.org/10.1016/j.yhbeh.2010.11.014.

Franzen, L., Cabugao, A., Grohmann, B., Elalouf, K., Johnson, A.P., Spitschan, M., 2022. Individual pupil size changes as a robust indicator of cognitive familiarity differences.

PLoS One 17 (1), e0262753. https://doi.org/10.1371/journal.pone.0262753.
Fullerton, J. N., Segre, E., De Maeyer, R. P. H., Maini, A. A. N., & Gilroy, D. W. (2016). Intravenous endotoxin challenge in Healthy Humans: an experimental platform to investigate and modulate systemic inflammation. J. Visual. Exp., (111). 10.3791/53913.
Geangu, E., Hauf, P., Bhardwaj, R., Bentz, W., Sirigu, A., 2011. Infant pupil diameter changes in response to others' positive and negative emotions. PLoS One 6 (11), e27132. https://doi.org/10.1371/journal.pone.0027132.
Glaholt, M.G., Wu, M.C., Reingold, E.M., 2009. Predicting preference from fixations. PsychNology Journal 7 (2), 141–158. https://doi.org/10.1037/e527342012-455.
Goodwin, R., Wiwattanapantuwong, J., Tuicomepee, A., Suttiwan, P., Watakakosol, R., 2020. Anxiety and public responses to covid-19: early data from Thailand. J. Psychiatr. Res. 129, 118–121. https://doi.org/10.1016/j.jpsychires.2020.06.026.
Gredebäck, G., Melinder, A., 2010. Infants' understanding of everyday social interactions: A dual process account. Cognition 114 (2), 197–206. https://doi.org/10.1016/j.cognition.2009.09.004.
Gul, P., Keesmekers, N., Elmas, P., Köse, F.E., Koskun, T., Wisman, A., Kupfer, T.R., 2021. Disease avoidance motives trade-off against social motives, especially mate-seeking, to predict social distancing: Evidence from the Covid-19 pandemic. Soc. Psychol. Personal. Sci. https://doi.org/10.1177/19485506211046462.
Hepach, R., Westermann, G., 2016. Pupillometry in infancy research. J. Cogn. Dev. 17 (3), 359–377. https://doi.org/10.1080/15248372.2015.1135801.
Hönekopp, J., 2006. Once more: is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. J. Exp. Psychol. Hum. Percept. Perform. 32 (2), 199. https://doi.org/10.1037/0096-1523.32.2.199.
Ho-Phuoc, T., Guyader, N., Guérin-Dugué, A., 2010. A functional and statistical bottom-up saliency model to reveal the relative contributions of low-level visual guiding factors. Cogn. Comput. 2 (4), 344–359. https://doi.org/10.1007/s12559-010-9078-8.
Hugenberg, K., Bodenhausen, G.V., 2003. Facing prejudice: Implicit prejudice and the perception of facial threat. Psychol. Sci. 14 (6), 640–643. https://doi.org/10.1046/j.0956-7976.2003.psci_1478.x.
Inagaki, T.K., Muscatell, K.A., Irwin, M.R., Moieni, M., Dutcher, J.M., Jevtic, I., Breen, E.C., Eisenberger, N.I., 2015. The role of the ventral striatum in inflammatory-induced approach toward support figures. Brain Behav. Immun. 44, 247–252. https://doi.org/10.1016/j.bbi.2014.10.006.
Jack, R.E., Schyns, P.G., 2015. The human face as a dynamic tool for social communication. Curr. Biol. 25 (14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052.
Jackson, I., Sirois, S., 2009. Infant cognition: Going full factorial with pupil dilation. Dev. Sci. 12 (4), 670–679. https://doi.org/10.1111/j.1467-7687.2008.00805.x.
Jones, A.L., Batres, C., Porcheron, A., Sweda, J.R., Morizot, F., Russell, R., 2018. Positive facial affect looks healthy. Vis. Cogn. 26 (1), 1–12. https://doi.org/10.1080/13506285.2017.1369202.
Jones, B.C., Perrett, D.I., Little, A.C., Boothroyd, L., Cornwell, R.E., Feinberg, D.R., Tiddeman, B.P., Whiten, S., Pitman, R.M., Hillier, S.G., et al., 2005. Menstrual cycle, pregnancy and oral contraceptive use alters attraction to apparent health in faces. Proc. R. Soc. B Biol. Sci. 272 (1561), 347–354. https://doi.org/10.1098/rspb.2004.2962.
Kalick, S.M., Zebrowitz, L.A., Langlois, J.H., Johnson, R.M., 1998. Does human facial attractiveness honestly advertise health? Longitudinal data on an evolutionary question. Psychol. Sci. 9 (1), 8–13. https://doi.org/10.1111/1467-9280.00002.
Karlsson, E.K., Kwiatkowski, D.P., Sabeti, P.C., 2014. Natural selection and infectious disease in human populations. Nat. Rev. Genet. 15 (6), 379–393. https://doi.org/10.1038/nrg3734.
Kikuchi, Y., Senju, A., Akechi, H., Tojo, Y., Osanai, H., Hasegawa, T., 2011. Atypical disengagement from faces and its modulation by the control of eye fixation in children with autism spectrum disorder. J. Autism Dev. Disord. 41 (5), 629–645. https://doi.org/10.1007/s10803-010-1082-z.
Kramer, P., Bressan, P., 2021. Infection threat shapes our social instincts. Behav. Ecol. Sociobiol. 75 (3), 47. https://doi.org/10.1007/s00265-021-02975-9.
Kret, M.E., Sjak-Shie, E.E., 2018. Preprocessing pupil size data: Guidelines and code. Behav. Res. Methods 51, 1336–1342. https://doi.org/10.3758/s13428-018-1075-y.
Kret, M.E., Roelofs, K., Stekelenburg, J., de Gelder, B., 2013. Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. Front. Hum. Neurosci. 7, 810. https://doi.org/10.3389/fnhum.2013.00810.
Kuraguchi, K., Kanari, K., 2021. Enlargement of female pupils when perceiving something cute. Sci. Rep. 11 (1), 23367. https://doi.org/10.1038/s41598-021-02852-5.
Kurzban, R., Leary, M.R., 2001. Evolutionary origins of stigmatization: the functions of social exclusion. Psychol. Bull. 127 (2), 187. https://doi.org/10.1037//0033-2909.127.2.187.
Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2017. LmerTest: Package: Tests in linear mixed effects models. J. Stat. Softw., 82(13). 10.18637/jss.v082.i13.
Langton, S.R., Law, A.S., Burton, A.M., Schweinberger, S.R., 2008. Attention capture by faces. Cognition 107 (1), 330–342. https://doi.org/10.1016/j.cognition.2007.07.012.
Lasselin, J., Sundelin, T., Wayne, P.M., Olsson, M.J., Göranson, S.P., Axelsson, J., Lekander, M., 2020. Biological motion during inflammation in humans. Brain Behav. Immun. 84, 147–153. https://doi.org/10.1016/j.bbi.2019.11.019.
Leder, H., Tinio, P.P., Fuchs, I.M., Bohrn, I., 2010. When attractiveness demands longer looks: The effects of situation and gender. Q. J. Exp. Psychol. 63 (9), 1858–1871. https://doi.org/10.1080/17470211003605142.
Lee, S.W., Schwarz, N., Taubman, D., Hou, M., 2010. Sneezing in times of a flu pandemic: public sneezing increases perception of unrelated risks and shifts preferences for federal spending. Psychol. Sci. 21 (3), 375–377. https://doi.org/10.1177/0956797609359876.
Lemay, Jr, E.P., Clark, M.S., Greenberg, A., 2010. What is beautiful is good because what is beautiful is desired: Physical attractiveness stereotyping as projection of interpersonal goals. Pers. Soc. Psychol. Bull. 36 (3), 339–353. https://doi.org/10.1177/0146167209359700.
Leschak, C.J., Hornstein, E.A., Haltom, K.E.B., Johnson, K.L., Breen, E.C., Irwin, M.R., Eisenberger, N.I., 2022. Ventromedial prefrontal cortex activity differentiates sick from healthy faces: Associations with inflammatory responses and disease avoidance motivation. Brain Behav. Immun. 100, 48–54. https://doi.org/10.1016/j.bbi.2021.11.011.
Liu, L., Johnson, H.L., Cousens, S., Perin, J., Scott, S., Lawn, J.E., Rudan, I., Campbell, H., Cibulskis, R., Li, M., et al., 2012. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet 379 (9832), 2151–2161. https://doi.org/10.1016/s0140-6736(12)60560-1.
Maas, C. J. M., Hox, J. J., 2005. Sufficient sample sizes for multilevel modeling. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 1, 86–92. https://doi.org/10.1027/1614-2241.1.3.86.
Makhanova, A., Shepherd, M.A., 2020. Behavioral immune system linked to responses to the threat of COVID-19. Pers. Individ. Differ. 167, 110221. https://doi.org/10.1016/j.paid.2020.110221.
Makowski, D., Wiernik, B. M., Patil, I., Lüdecke, D., Ben-Shachar, M. S., White, M., & Rabe, M. M. 2022. Correlation: Methods for Correlation Analysis (0.8.3). https://CRAN.R-project.org/package=correlation.
Manzoor, S., Latifi, S. 2021. A health detection model based on facial data. In ITNG 2021 18th International Conference on Information Technology-New Generations (pp. 463-468). Cham: Springer International Publishing. 10.1007/978-3-030-70416-2_60.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164 (1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.
Michalak, N.M., Sng, O., Wang, I.M., Ackerman, J., 2020. Sounds of sickness: can people identify infectious disease using sounds of coughs and sneezes? Proc. R. Soc. B 287 (1928), 20200944. https://doi.org/10.1098/rspb.2020.0944.
Moshkin, M., Litvinova, N., Litvinova, E.A., Bedareva, A., Lutsyuk, A., Gerlinskaya, L., 2012. Scent recognition of infected status in humans. J. Sex. Med. 9 (12), 3211–3218. https://doi.org/10.1111/j.1743-6109.2011.02562.x.
Muscatell, K.A., Inagaki, T.K., 2021. Beyond social withdrawal: New perspectives on the effects of inflammation on social behavior. Brain Behav. Immunity-Health 16, 100302. https://doi.org/10.1016/j.bbih.2021.100302.
Nabeela, F., Azizullah, A., Bibi, R., Uzma, S., Murad, W., Shakir, S.K., Ullah, W., Qasim, M., Häder, D.P., 2014. Microbial contamination of drinking water in Pakistan—a review. Environ. Sci. Pollut. Res. 21 (24), 13929–13942. https://doi.org/10.1007/s11356-014-3348-z.
Nikitin, J., Freund, A.M., 2019. The motivational power of the happy face. Brain Sci. 9 (1), 6. https://doi.org/10.3390/brainsci9010006.
Nunner, H., Buskens, V., Kretzschmar, M., 2021. A model for the co-evolution of dynamic social networks and infectious disease dynamics. Comput. Soc. Netw. 8 (1), 1–33. https://doi.org/10.1186/s40649-021-00098-9.
Olsson, M.J., Lundström, J.N., Kimball, B.A., Gordon, A.R., Karshikoff, B., Hosseini, N., Sorjonen, K., Olgart Höglund, C., Solares, C., Soop, A., Axelsson, J., Lekander, M., 2014. The scent of disease: human body odor contains an early chemosensory cue of sickness. Psychol. Sci. 25 (3), 817–823.
Powell, H., Smart, J., Wood, L.G., Grissell, T., Shafren, D.R., Hensley, M.J., Gibson, P.G., Mant, D., 2008. Validity of the Common Cold Questionnaire (CCQ) in asthma exacerbations. PLoS One 3 (3), e1802. https://doi.org/10.1371/journal.pone.0001802.
Regenbogen, C., Axelsson, J., Lasselin, J., Porada, D.K., Sundelin, T., Peter, M.G., Lekander, M., Lundström, J.N., Olsson, M.J., 2017. Behavioral and neural correlates to multisensory detection of sick humans. Proc. Natl. Acad. Sci. U.S.A. 114 (24), 6400–6405. https://doi.org/10.1073/pnas.1617357114.
Roesmann, K., Wiens, N., Winker, C., Rehbein, M.A., Wessing, I., Junghoefer, M., 2020. Fear generalization of implicit conditioned facial features–Behavioral and magnetoencephalographic correlates. Neuroimage 205, 116302. https://doi.org/10.1016/j.neuroimage.2019.116302.
Sarolidou, G., Axelsson, J., Sundelin, T., Lasselin, J., Regenbogen, C., Sorjonen, K., Lundström, J.N., Lekander, M., Olsson, M.J., 2019. Emotional expressions of the sick face. Brain Behav. Immun. 80, 286–291. https://doi.org/10.1016/j.bbi.2019.04.003.
Sarolidou, G., Axelsson, J., Kimball, B.A., Sundelin, T., Regenbogen, C., Lundström, J.N., Olsson, M.J., 2020a. People expressing olfactory and visual cues of disease are less liked. Philos. Trans. R. Soc. B 375 (1800), 20190272. https://doi.org/10.1098/rstb.2019.0272.
Sarolidou, G., Tognetti, A., Lasselin, J., Regenbogen, C., Lundström, J.N., Kimball, B.A., Garke, M., Lekander, M., Axelsson, J., Olsson, M.J., 2020b. Olfactory communication of sickness cues in respiratory infection. Front. Psychol. 11. https://doi.org/10.3389/fpsyg.2020.01004.
Schaller, M., Murray, D.R., Bangerter, A., 2015. Implications of the behavioural immune system for social behaviour and human health in the modern world. Philos. Trans. R. Soc., B 370 (1669), 20140105. https://doi.org/10.1098/rstb.2014.0105.
Schaller, M., Park, J.H., 2011. The behavioral immune system (and why it matters). Curr. Dir. Psychol. Sci. 20 (2), 99–103. https://doi.org/10.1177/0963721411402596.
Schaller, M., Park, J., Faulkner, J., 2003. Prehistoric dangers and contemporary prejudices. Eur. Rev. Soc. Psychol. 14 (1), 105–137. https://doi.org/10.1080/10463280340000036.
Schedlowski, M., Engler, H., Grigoleit, J.S., 2014. Endotoxin-induced experimental systemic inflammation in humans: a model to disentangle immune-to-brain communication. Brain Behav. Immun. 35, 1–8. https://doi.org/10.1016/j.bbi.2013.09.015.
Schrock, J.M., Snodgrass, J.J., Sugiyama, L.S., 2020. Lassitude: The emotion of being sick.

Evol. Hum. Behav. 41 (1), 44–57. https://doi.org/10.1016/j.evolhumbehav.2019.09.002.
Shapouri, S., Martin, L. L. 2021. Snakes vs. guns: A systematic review of comparisons between phylogenetic and ontogenetic threats. Adapt. Hum. Behav. Physiol., 131-155. 10.1007/s40750-021-00181-5.
Shirasu, M., Touhara, K., 2011. The scent of disease: Volatile organic compounds of the human body related to disease and disorder. J. Biochem. 150 (3), 257–266. https://doi.org/10.1093/jb/mvr090.
Smillie, L.D., Jackson, C.J., 2005. The appetitive motivation scale and other BAS measures in the prediction of approach and active avoidance. Pers. Individ. Differ. 38 (4), 981–994. https://doi.org/10.1016/j.paid.2004.09.013.
Sui, J., Liu, C.H., 2009. Can beauty be ignored? Effects of facial attractiveness on covert attention. Psychon. Bull. Rev. 16 (2), 276–281. https://doi.org/10.3758/PBR.16.2.276.
Sundelin, T., Karshikoff, B., Axelsson, E., Höglund, C.O., Lekander, M., Axelsson, J., 2015. Sick man walking: Perception of health status from body motion. Brain Behav. Immun. 48, 53–56. https://doi.org/10.1016/j.bbi.2015.03.007.
Sundelin, T., Lekander, M., Sorjonen, K., Axelsson, J., 2017. Negative effects of restricted sleep on facial appearance and social appeal. R. Soc. Open Sci. 4 (5), 160918. https://doi.org/10.1098/rsos.160918.
Tskhay, K.O., Wilson, J.P., Rule, N.O., 2016. People use psychological cues to detect physical disease from faces. Pers. Soc. Psychol. Bull. 42 (10), 1309–1320. https://doi.org/10.1177/0146167216656357.
Tybur, J.M., Jones, B.C., DeBruine, L.M., Ackerman, J.M., Fasolt, V., 2020. Preregistered direct replication of "sick body, vigilant mind: The biological immune system activates the behavioral immune system". Psychol. Sci. 31 (11), 1461–1469. https://doi.org/10.1177/0956797620955209.
Tybur, J. M., Lieberman, D., Kurzban, R., DeScioli, P., 2013. Disgust: evolved function and structure. Psychological review 120 (1), 65. https://doi.org/10.1037/a0030778.
van Leeuwen, F., Petersen, M.B., 2018. The behavioral immune system is designed to avoid infected individuals, not outgroups. Evol. Hum. Behav. 39 (2), 226–234. https://doi.org/10.1016/j.evolhumbehav.2017.12.003.
van Reekum, C.M., Johnstone, T., Urry, H.L., Thurow, M.E., Schaefer, H.S., Alexander, A.L., Davidson, R.J., 2007. Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. Neuroimage 36 (3), 1041–1055. https://doi.org/10.1016/j.neuroimage.2007.03.052.
Walther, D., Koch, C., 2006. Modeling attention to salient proto-objects. Neural Netw. 19 (9), 1395–1407. https://doi.org/10.1016/j.neunet.2006.10.001.
Wang, C.A., Baird, T., Huang, J., Coutinho, J.D., Brien, D.C., Munoz, D.P., 2018. Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. Front. Neurol. 9, 1029. https://doi.org/10.3389/fneur.2018.01029.
Wardle, S.G., Taubert, J., Teichmann, L., Baker, C.I., 2020. Rapid and dynamic processing of face pareidolia in the human brain. Nat. Commun. 11 (1), 4518. https://doi.org/10.1038/s41467-020-18325-8.
Waynforth, D., 2001. Mate choice trade-offs and women's preference for physically attractive men. Hum. Nat. 12 (3), 207–219. https://doi.org/10.1007/s12110-001-1007-9.
Westfall, J., Kenny, D.A., Judd, C.M., 2014. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. J. Exp. Psychol. Gen. 143 (5), 2020–2045. https://doi.org/10.1037/xge0000014.
Widen, S.C., Pochedly, J.T., Pieloch, K., Russell, J.A., 2013. Introducing the sick face. Motiv. Emot. 37, 550–557. https://doi.org/10.1007/s11031-013-9353-6.
Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., 937 Tanaka, J. W., 2010. Controlling low-level image properties: The SHINE toolbox. Behav. Res. Methods 42 (3), 671–684. https://doi.org/10.3758/BRM.42.3.671.
World Health Organization. 2020, December 9. The top 10 causes of death. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.