Motivation and Emotion (2024) 48:278–294
https://doi.org/10.1007/s11031-024-10076-z

ORIGINAL PAPER

A preliminary characterization of the psychometric properties and generalizability of a novel social approach-avoidance paradigm

Travis C. Evans[1,2,6] · Josie Carlson[3] · Agnieszka Zuberer[1,8] · Regan Fry[1] · Sam Agnoli[1] · Jennifer C. Britton[4] · Joseph DeGutis[1,5] · Michael Esterman[1,2,6,7]

Accepted: 4 May 2024 / Published online: 6 July 2024

Abstract
Social behaviors are guided in part by motivational and emotional responses to affective facial expressions. In daily life, facial expressions communicate varying degrees social reward signals (happiness), social threat signals (anger), or social reward-threat conflict signals (co-occurring happiness and anger). Thus, motivational and emotional responses must be sensitive to variations in social signal intensity to effectively guide social behavior. We recently developed a novel social approach-avoidance paradigm (SAAP), which uses morphed facial expressions to assess sensitivity to linear increases in social reward and/or social threat intensity. Prior to large-scale studies validating the test quality of the SAAP, however, it is necessary to first establish the psychometric properties and generalizability of these sensitivity metrics. In Study 1, we independently replicated SAAP task effects and demonstrated that motivational and emotional sensitivity measures exhibit strong psychometric properties and robust individual variability. In Study 2, we demonstrated that more complex social judgements (e.g., trustworthiness) are also sensitive to linear increases in social signal intensity, which differs across judgements. Although future research in larger samples will be needed to establish the test quality of the SAAP, these preliminary findings suggest that the SAAP exhibits adequate psychometric properties to justify this type of large-scale individual differences research.

Keywords Social · Approach · Avoidance · Conflict · Psychometric

✉ Travis C. Evans
tevans1@bu.edu

[1] Boston Attention and Learning Lab (BALLAB), VA Boston Healthcare System, Boston, U.S.A.

[2] Department of Psychiatry, Boston University Chobanian and Avedisian School of Medicine, Boston, U.S.A.

[3] Department of Neuroscience, Boston University, Boston, U.S.A.

[4] Department of Psychology, University of Miami, Miami, U.S.A.

[5] Department of Psychiatry, Harvard Medical School, Boston, U.S.A.

[6] National Center for PTSD, VA Boston Healthcare System, Boston, U.S.A.

[7] Neuroimaging Research for Veterans (NeRVe) Center, VA Boston Healthcare System, Boston, U.S.A.

[8] Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, United States

General introduction

Social behaviors are influenced by affective facial expressions that can signal opportunities for social affiliation and social rejection (Ambadar et al., 2005; Barrett et al., 2019; Frith, 2009; Rilling & Sanfey, 2011). Specifically, happy facial expressions are typically perceived as social reward signals that communicate an opportunity for social affiliation, whereas angry facial expressions are typically perceived as social threat signals that communicate an opportunity for social rejection (Chen & Jack, 2017; Nikitin & Freund, 2019; Roelofs et al., 2010; Seidel et al., 2010; Strack & Deutsch, 2004; Tamir & Hughes, 2018). As a result, happy facial expressions typically elicit approach motivational responses that facilitate social interactions, whereas angry facial expressions typically elicit avoidance motivational responses that disrupt or prevent social interactions (Marsh et al., 2005; Radke et al., 2018; Renard et al., 2017; Seidel et al., 2010; Stins et al., 2011; Vrana & Gross, 2004). In this manner, social behavior is strongly influenced by motivational and emotional responses to affective facial expressions, which are typically measured using experimental paradigms.

To date, experimental paradigms have largely examined motivational and emotional responses to affective facial expressions that display highly intense social reward signals (100%$_{Happy}$) or social threat signals (e.g., 100%$_{Angry}$). However, intense facial expressions rarely occur in daily life or in social contexts (Carroll & Russell, 1997). Instead, spontaneous facial expressions of emotion in commonly exhibit more subtle emotional signals or blends of multiple emotional signals (Scherer & Ceschi, 2000; Scherer & Ellgring, 2007). For example, affective facial expressions in social interactions tend to be more ambiguous due communicating varying intensities of social reward signals (e.g., 50%$_{Happy}$) or social threat signals (e.g., 50%Angry; Barrett et al., 2019; Beevers et al., 2009; Matsumoto & Hwang, 2014). Additionally, facial expressions also exhibit co-occurring social reward signals and social threat signals, which is facilitated by the strong independence of facial musculature along the vertical face axis (i.e., eyes and mouth) and frequently occurs within spontaneous facial expressions of emotion (Du & Martinez, 2015; Du et al., 2014; Ross et al., 2016). Importantly, individuals are also most accurate in recognizing emotions conveyed by blended facial expressions, which is presumably due to greater familiarity with blended facial expressions in daily life (Calvo et al., 2014; LaPlante & Ambady, 2000). Therefore, traditional experimental paradigms fail to capture motivational and emotional responses to subtle degrees of social reward signals and social threat signals that are more characteristic of facial expression during social interactions, which is necessary to capture ecologically valid measures that map on to social behavior (Carroll & Russell, 1997; Matsumoto & Hwang, 2014).

In addition to lacking ecological validity, measuring motivational and emotional responses to intense affective facial expressions may obfuscate individual differences due to ceiling effects in these responses. Consistent with this view, previous research demonstrates that individual differences may be greatest when facial expressions are ambiguous due to varying degrees of social reward and/or social threat signals (Evans et al., 2022; Gutiérrez-García & Calvo, 2014, 2016; Staugaard, 2010). Among ambiguous facial expressions, interindividual variability may be greatest when social reward signals and social threat signals simultaneously *co-occur* to communicate social reward-threat conflict signals, which activates competing motivations to approach *and* avoid another individual (Evans et al., 2022; Gutiérrez-García & Calvo, 2014, 2016). Along with the need to improve ecological validity, these findings underscore the importance of developing experimental paradigms that reliably measure the sensitivity of motivational and emotional sensitivity to subtle *variations* in social signal intensity.

To address this issue, we recently developed a novel social approach-avoidance paradigm (SAAP), which measures the sensitivity of approach-avoidance motivation and perceived emotion as a function of linear increases in social signal intensity (Evans & Britton, 2020; Evans et al., 2022). In the SAAP, affective facial expressions are linearly interpolated in 25% increments to create morphed facial expressions that parametrically modulate the signal intensity of social reward, social threat, or social reward-threat conflict (see Fig. 1). Our previous research using the SAAP demonstrates linear increases in both approach motivation and perceived happiness as a function of social reward intensity, whereas individuals exhibit linear increases in both avoidance motivation and perceived anger as a function of social threat intensity (Evans & Britton, 2020; Evans et al., 2022). As social reward increases along with *co-occurring* decreases in social threat (i.e., social reward-threat conflict), individuals exhibit linear increases in approach motivation and perceived happiness that *co-occur* with linear decreases in avoidance motivation and perceived anger (Evans & Britton, 2020; Evans et al., 2022). Importantly, our previous research using the SAAP demonstrates that social avoidance behavior is associated with dysregulated sensitivity of approach-avoidance motivation to more ambiguous social reward/social threat, but not in response to intense social reward or social threat (Evans et al., 2022). Demonstrating the utility of studying social reward-threat conflict in particular, social avoidance behavior was most strongly associated with weaker approach motivation sensitivity and stronger avoidance motivation sensitivity to *co-occurring* changes in social reward and social threat intensity, but not varying degrees of social reward intensity or social threat intensity in isolation (Evans et al., 2022). Together, these initial results suggest that the SAAP effectively measures motivational sensitivity to linear changes in social signal intensity, which demonstrates associations with maladaptive social behavior that are highly relevant to psychopathology.

Prior to establishing if the SAAP is suitable for large-scale individual differences research, however, it is first necessary to ascertain if the SAAP produces replicable task effects with adequate psychometric properties (Cooper et al., 2017; Fröhner et al., 2019; Hayden, 2022; Parsons et al., 2019). First, to reduce the risk of false positive findings, motivational and emotional sensitivity in the SAAP should exhibit robust effect sizes that replicate across independent samples (Miller & Ulrich, 2013; Nosek et al., 2022). Second, to reliably and economically detect associations with other measures of interest, SAAP task measures should demonstrate adequate internal consistency with relatively few task trials (Evans & Britton, 2018; Evans et al., 2018; Goodhew & Edwards, 2019; Infantolino et al., 2018; Nunnally, 1967). Third, to characterize more trait-like individual differences and reliably quantify the effects of experimental manipulations, SAAP task measures should demonstrate test–retest

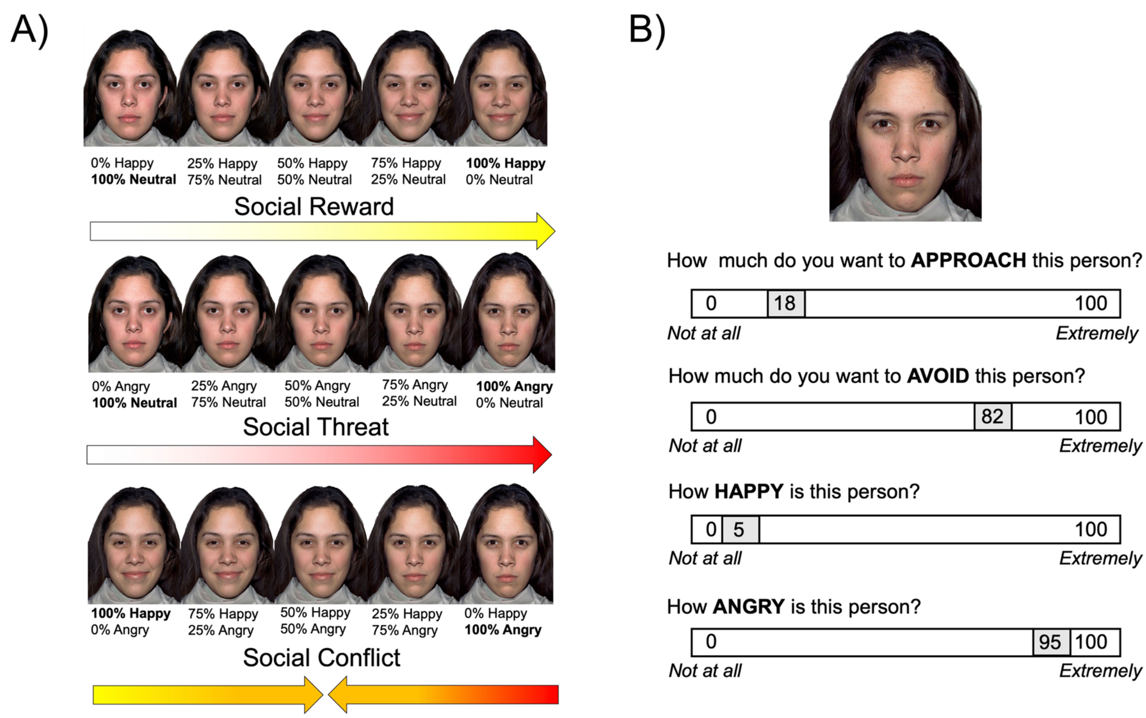

Fig. 1 The Social Approach-Avoidance Paradigm. *Legend:* **A**) Morphed facial expressions were created by visually blending neutral, happy, and angry facial expressions to parametrically modulate social reward signals (*Top*: $0\%_{\text{Happy}}$, $25\%_{\text{Happy}}$, $50\%_{\text{Happy}}$, $75\%_{\text{Happy}}$, or $100\%_{\text{Happy}}$), social threat signals (*Middle*: $0\%_{\text{Angry}}$, $25\%_{\text{Angry}}$, $50\%_{\text{Angry}}$, $75\%_{\text{Angry}}$, or 100%), and social reward-threat conflict (*Bottom*: $100\%_{\text{Happy}}+0\%_{\text{Angry}}$, $75\%_{\text{Happy}}+25\%_{\text{Angry}}$, $50\%_{\text{Happy}}+50\%_{\text{Angry}}$, $25\%_{\text{Happy}}+75\%_{\text{Angry}}$, or $0\%_{\text{Happy}}+100\%_{\text{Angry}}$). (**B**) In the version of the Social Approach-Avoidance Paradigm (SAAP) used in the current study, a facial expression is presented along with a visual rating scale ranging from 0 (*Not at All*) to 100 (*Extremely*). Using this visual rating scale, participants rated the degree to which they would feel motivated to approach or avoid the individual as well as how happy or angry they perceived the facial expression

reliability across time (McCrae et al., 2011; Polit, 2014). Fourth, to facilitate analyses of covariance in individual differences research, SAAP task measures should robustly differ *between* individuals (Goodwin & Leech, 2006; Hedge et al., 2018; Martinez et al., 2020). Addressing these criteria is an important precursor to conducting more large-scale individual differences research that can assess convergent and discriminant validity and establishing normative test values for the SAAP. Thus, prior to conducting large-scale individual differences research using the SAAP, it is necessary to first establish that SAAP task measures exhibit reliability and replicability, internal consistency, test–retest reliability, and between-subject variability.

In addition to preliminarily characterizing the psychometric properties of SAAP task measures, it is also important to assess the generalizability of the SAAP beyond basic motivation and emotion perception. Specifically, social signals within facial expressions are used to generate a variety of complex interpersonal judgements, which contribute to social behavior across a wide variety of contexts (for a review, see Todorov et al., 2015). For example, individuals utilize varying degrees of happiness and/or anger within facial expressions to generate more complex interpersonal judgements about another person's trustworthiness or potential for aggression (for a review, see Todorov, 2008). Thus, it is important to determine if SAAP task effects generalize beyond basic motivational and emotion perceptual processes to more complex interpersonal judgements that also play an important role in social behavior.

Objectives of the current studies

To address these issues, we conducted two experiments that respectively aimed to characterize the psychometric properties (Study 1) and generalizability (Study 2) of sensitivity to linear changes in social reward and/or social threat in the SAAP. In Study 1, we administered the SAAP to two independent samples of adults who completed the task either in the laboratory or online. Using these samples, we characterized the psychometric properties of SAAP task measures as well as if SAAP task measures differed between laboratory task administration and online task administration. In Study 2, we administered the SAAP using a variety of more complex interpersonal judgements in addition to basic motivation and emotion perception ratings. Thus, the primary aim of Study 2 was to determine if more complex interpersonal judgements also exhibit sensitivity to linear changes in social reward/social threat in the SAAP as well as the degree to which sensitivity differed among interpersonal judgements.

Study 1

Transparency and openness

Study 1 was not preregistered prior to data collection and analysis. De-identified data for the online sample for Study 1 and all analysis syntax for Study 1 is available in an open-access data repository (https://osf.io/yh42j/ bb2c5b-f777144b39a75c8a6ed158007d). At the time of laboratory data collection, our IRB protocol did not include a provision for making those data publicly available. Morphed facial expression stimuli can be made freely available to researchers who receive written permission to access the NimStim set (Tottenham et al., 2009).

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in Study 1. All participants provided written informed consent and study procedures were conducted in accordance with local IRB guidelines at the University of Miami (Protocol: #20-120-901). All study procedures complied with the Helsinki Declaration as revised in 2008.

Methods

Participants

Laboratory sample

As reported in our previous work (Evans & Britton, 2020), the laboratory sample was comprised of 57 adults (64.90% Female; Age: $M=20.59$ years, $SD=3.66$ years; 54.40% Caucasian, 15.80% Black, 21.10% Asian, 3.50% Pacific Islander, and 5.30% multi-racial) recruited from the University of Miami and surrounding Miami community. Sample size for the laboratory sample was determined based on effect sizes reported in a previous study on automatic approach-avoidance actions (Kuckertz et al., 2017). In exchange for participation, participants received either monetary payment ($10 per hour) and/or course credit.

Online sample

The online sample was comprised of 60 undergraduate students (53.30% Female; Age: $M=19.37$ years, $SD=1.48$ years; 68.30% Caucasian, 10.00% Black, 16.70% Asian, 1.70% Pacific Islander, and 3.30% multi-racial) recruited from the University of Miami. From this larger sample, 31 participants (54.80% Female; Age: $M=19.45$, $SD=1.52$ years; 64.50% Caucasian, 9.70% Black, and 25.80% Asian) completed the SAAP at a second online session approximately 2 weeks later ($M=12.50$ days, $SD=6.78$ days). Sample size for the online sample was selected to match our laboratory sample. Participants who completed both online sessions did not significantly differ in demographic characteristics or SAAP task effects compared to participants who only completed the first online session (all $ps>0.16$). In exchange for participation, participants received course credit.

Stimuli and task

Morphed facial expressions

Morphed facial expressions were created using Morpheus software, which served to parametrically modulate social signals in 25% increments of intensity (e.g., $0\%_{\text{Happy}}$, $25\%_{\text{Happy}}$, $50\%_{\text{Happy}}$, $75\%_{\text{Happy}}$, and $100\%_{\text{Happy}}$). Specifically, we linearly interpolated stereotypic happy, angry, and neutral facial expressions (i.e., $100\%_{\text{Happy}}$, $100\%_{\text{Angry}}$, and $100\%_{\text{Neutral}}$) from the same actor to parametrically modulate social reward, social threat, and social reward-threat conflict signals. For example, we linearly interpolated $100\%_{\text{Neutral}}$ and $100\%_{\text{Happy}}$ facial expressions from the same actor in 25% increments to parametrically modulate social reward signals (i.e., $0\%_{\text{Happy}}$, $25\%_{\text{Happy}}$, $50\%_{\text{Happy}}$, $75\%_{\text{Happy}}$, and $100\%_{\text{Happy}}$). Using this method, we created morphed facial expressions that parametrically varied in either social reward signals (e.g., $50\%_{\text{Happy}}$), social threat signals (e.g., $50\%_{\text{Angry}}$), or social reward-threat conflict signals (e.g., $50\%_{\text{Happy}}+50\%_{\text{Angry}}$; see Fig. 1A). Morphed facial expressions were generated using six male actors and six female actors from the NimStim stimulus set (Tottenham et al., 2009).

Social Approach-Avoidance Paradigm (SAAP)

In the SAAP, participants provide subjective ratings in response to each facial expression using 101-point visual scales (0 = *Not at all*; 100 = *Extremely*), which dynamically updated based on slider movement in 1% increments (see Fig. 1B). In response to each randomly presented facial expression, participants rated their motivation to approach and avoid the individual (approach-avoidance motivation) as well as perceived happiness and anger (emotional perception). After completing ratings for a facial expression, participants advanced to the next randomly presented facial expression in a self-paced manner. Facial expressions were presented in a fully randomized order across participants. For both the lab sample and the online sample, the SAAP was programmed and administered using Qualtrics software (Qualtrics, Provo, UT).

In the SAAP, morphed facial expressions were created using faces from six male actors and six female actors. Including the unmorphed facial expressions ($100\%_{Neutral}$, $100\%_{Happy}$, and $100\%_{Angry}$), there are a total of 12 facial expressions for each actor (see Fig. 1A). To economize task administration, participants rated morphed facial expressions from a randomized selection of three male actors and three female actors. For these six randomly selected actors, participants completed all ratings for each of the 12 facial expressions (e.g., $50\%_{Happy}$). For example, participants provided one approach rating for the $50\%_{Happy}$ facial expression of each actor, which resulted in 6 separate approach ratings for the $50\%_{Happy}$ facial expression. In this manner, participants completed all ratings for each of the 12 facial expressions across 6 randomly selected actors, which resulted in 72 ratings (e.g., 12 facial expressions * 6 face identities = 72 approach ratings). In total, participants completed 288 ratings during the SAAP (i.e., 72 approach ratings, 72 avoid ratings, 72 happiness ratings, and 72 angry ratings).

Data analytic strategy

All analyses were conducted using either SPSS software ver. 24.0 (IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM) or R software (R Core Team, 2021). For GLMM analyses, we utilized R library packages including lme4 and sjPlot (Bates et al., 2015; Lüdecke, 2018).

Replication of SAAP sensitivity measures

As described previously, the SAAP presents morphed facial expressions that linearly vary in degree of social reward, social threat, or co-occurring social reward and social threat. Therefore, we primarily characterized SAAP task effects using Repeated Measures ANOVA (RM-ANOVA) models that assessed *linear* increases or decreases in subjective ratings as a function of social signal intensity. In line with our previous research, we submitted subjective ratings to polynomial RM-ANOVA models that tested a 2 (Rating) × Morph (Linear Contrast) interaction effect.[1] However, this analytic approach is potentially limited because it assumes a linear response pattern, which will fail to characterize any *non-linear* response patterns. Thus, we also conducted secondary analyses with a RM-ANOVA model that characterizes SAAP task effects without assuming a linear response pattern (i.e., 2 (Rating) × Morph (0% vs. 25% vs. 50% vs. 75% vs. 100%).

Comparison of SAAP sensitivity measures between samples

Even in the case of a successful independent replication, it nevertheless remains possible that SAAP task effects systematically differ between laboratory and online task administration, which is an important consideration for study design. Therefore, we also directly compared SAAP task measures between the laboratory sample and online sample by testing for 2 (Sample: Lab vs. Online) × 2 (Rating) × Morph interaction effects.

Internal consistency of SAAP sensitivity measures

To characterize the internal consistency of SAAP task effects, we computed SAAP task effects using several computation methods aligned with the previously described group-level analyses. First, we computed linear slope estimates by multiplying subjective ratings with the same linear polynomial contrast vector used in our primary RM-ANOVA models (-2, -1, 0 , 1, 2), which is identical to conducting a linear regression to estimate the best fitting linear slope. Following this computation, scalar values were summed to create a single continuous score that indexes the strength (i.e., slope magnitude) and direction (i.e., positive or negative) of rating slopes as a function of social reward, social threat, or social reward-threat conflict. Additionally, we also computed the internal consistency for the *difference* in slopes for approach-avoidance motivation sensitivity (i.e., Approach slope – Avoid slope) and emotion perceptual sensitivity (i.e., Happiness slope – Anger slope). These differences scores were used to characterize the internal consistency of the 2 (Rating) × Linear (Morph) *interaction* effects in the SAAP.

However, it is important to note that linear slope estimates will not capture non-linear response patterns in the SAAP due to employing polynomial linear contrasts that mathematically exclude responses to the 50% morphed stimuli (e.g., $\left(\left[0\%_{Happy}\right] * -2\right) + \left(\left[25\%_{Happy}\right] * -1\right) + \left(\left[50\%_{Happy}\right] * 0\right) + \left(\left[75\%_{Happy}\right] * 1\right) + \left(\left[100\%_{Happy}\right] * 2\right)$). Therefore, we also quantified internal consistency estimates using adjacent contrast scores (i.e., 0% vs. 25% vs. 50% vs. 75% vs. 100%), which equally weights all stimuli in alignment with our secondary RM-ANOVA models. Finally, we also computed internal consistency estimates for each morphed facial expression separately to ascertain the degree to which these measures may support other, more complex computational approaches.

Given limitations inherent to Cronbach's alpha (Cho & Kim, 2015; Revelle & Zinbarg, 2009), we computed Mcdonald's omega to characterize the internal consistency

[1] Overall, we observed that a linear slope term provided a strong fit to participant response patterns in the SAAP (Adjusted Linear $R^2 = 0.47$), whereas including a non-linear slope term did not substantively increase the variance explained for participant response patterns (Adjusted Quadratic $R^2 = 0.50$). Together, these results suggest that 1) Participant response patterns in the SAAP are predominantly linear in trajectory and 2). Responses to intermediate stimulus values (e.g., $50\%_{Happy}$) do not substantively deviate from these largely linear response patterns.

estimates for the slope values, contrast scores, and individual morph ratings. Although Mcdonald's omega can be interpreted in a manner similar to Cronbach's alpha, we also report estimates obtained using Cronbach's alpha to be comprehensive (see Supplemental Information).

Test–Retest reliability of SAAP sensitivity measures

To characterize test–retest reliability, we computed intraclass correlation coefficients (ICCs) between average slope estimates measured at Time 1 and average slope estimates measured at Time 2. We also characterized test–retest reliability for the aforementioned adjacent contrast scores in which each morphed facial expression is equally weighted. Finally, we also characterized test–retest reliability for each morphed intensity value separately. For all estimates, we calculated ICCs using a 2-way random effects model to account for potential heterogeneity in temporal stability across participants. To be conservative, we interpreted ICC values in a manner similar to internal consistency estimates (for a review, see Polit, 2014).

Individual differences in SAAP sensitivity measures

To characterize individual differences in SAAP task measures, we utilized generalized linear mixed models (GLMMs) that formally tested the degree of between-subject variability (Kliegl et al., 2011; Magezi, 2015). In the current study, random intercept terms modeled between-subject variability in the *magnitude* of subjective ratings, whereas random slope terms modeled between-subject variability in the *sensitivity* of subjective ratings to varying social signal intensity. To quantify between-subject variability in these GLMMs, we used likelihood-ratio tests to separately compare model fit between a null distribution model that constrained between-subject variability (i.e., no random effects terms) against a random intercept model and a random slope model (Bolker et al., 2009; Dingemanse & Dochtermann, 2013; Kliegl et al., 2011). To facilitate model convergence, we utilized bound optimization by quadratic approximation (BOBYQA) with a set maximum of 200,000 iterations. To compute GLMM p-values and degrees of freedom, we used maximum likelihood estimates in conjunction with the Satterthwaite approximation (Luke, 2017).

Results

Replication of SAAP sensitivity measures

In brief, we observed robust modulation of both approach-avoidance motivational sensitivity and emotion perceptual sensitivity (see statistical models below and Fig. 2). Approach and happiness ratings robustly increased as a function of social reward, whereas avoidance and anger ratings increased as a function of social threat. As a function of *co-occurring* social reward and social threat (social reward-threat conflict), both samples exhibited robust changes in approach and happiness ratings that *co-occurred* with opposing changes in avoidance and anger ratings, respectively.

For approach-avoidance motivation, we observed significant Rating × Morph (Linear) interaction effects for the social reward model (<u>Lab</u>: $F_{(1, 56)} = 152.45$, $p < 0.001$; <u>Online</u>: $F_{(1, 59)} = 191.09$, $p < 0.001$), social threat model (<u>Lab</u>: $F_{(1, 56)} = 110.83$, $p < 0.001$; <u>Online</u>: $F_{(1, 59)} = 201.75$, $p < 0.001$), and social reward-threat conflict model (<u>Lab</u>: $F_{(1, 56)} = 215.89$, $p < 0.001$; <u>Online</u>: $F_{(1, 59)} = 275.85$, $p < 0.001$). For all approach-avoidance models, we also observed robust and significant SAAP task effects when using RM-ANOVA models that equally weight all facial expressions without assumine a linear response pattern (all Fs > 86.06, all ps < 0.001).

For emotion perception, we similarly observed significant Rating × Morph (Linear) interaction effects for the social reward model (<u>Lab</u>: $F_{(1, 56)} = 564.81$, $p < 0.001$; <u>Online</u>: $F_{(1, 59)} = 590.64$, $p < 0.001$), social threat model (<u>Lab</u>: $F_{(1, 56)} = 331.71$, $p < 0.001$; <u>Online</u>: $F_{(1, 59)} = 502.27$, $p < 0.001$), and social reward-threat conflict model (<u>Lab</u>: $F_{(1, 56)} = 808.61$, $p < 0.001$; <u>Online</u>: $F_{(1, 59)} = 826.08$, $p < 0.001$). For all emotion perception models, we also observed significant, robust SAAP task effects when using RM-ANOVA models that equally weight all facial expressions and do not assume a linear response pattern (all Fs > 229.42, all ps < 0.001).

Comparison of SAAP sensitivity measures between samples

As quantified by Sample × Rating × Morph (Linear) interaction effects, the lab and online samples did not exhibit statistically significant differences in either approach-avoidance motivational sensitivity or emotional perception sensitivity (see Fig. 1C). Specifically, the lab sample and online sample did not exhibit statistically significant differences in approach-avoidance motivation sensitivity in the social reward model ($F_{(1, 115)} = 0.01$, $p = 0.93$), social threat model ($F_{(1, 115)} = 0.06$, $p = 0.80$), or social reward-threat conflict model ($F_{(1, 115)} = 0.13$, $p = 0.72$; see Fig. 1). Similarly, the lab sample and online sample did not exhibit statistically significant differences in emotion perception sensitivity in the social reward model ($F_{(1, 115)} = 1.22$, $p = 0.27$), social threat model ($F_{(1, 115)} = 0.49$, $p = 0.49$), or social reward-threat conflict model ($F_{(1, 115)} = 1.54$, $p = 0.22$). For all models, we also did not observe statistically significant differences in SAAP task effects between samples when using RM-ANOVA models that equally weighted all morphed facial expressions (all Fs < 1.61, all ps > 0.21).

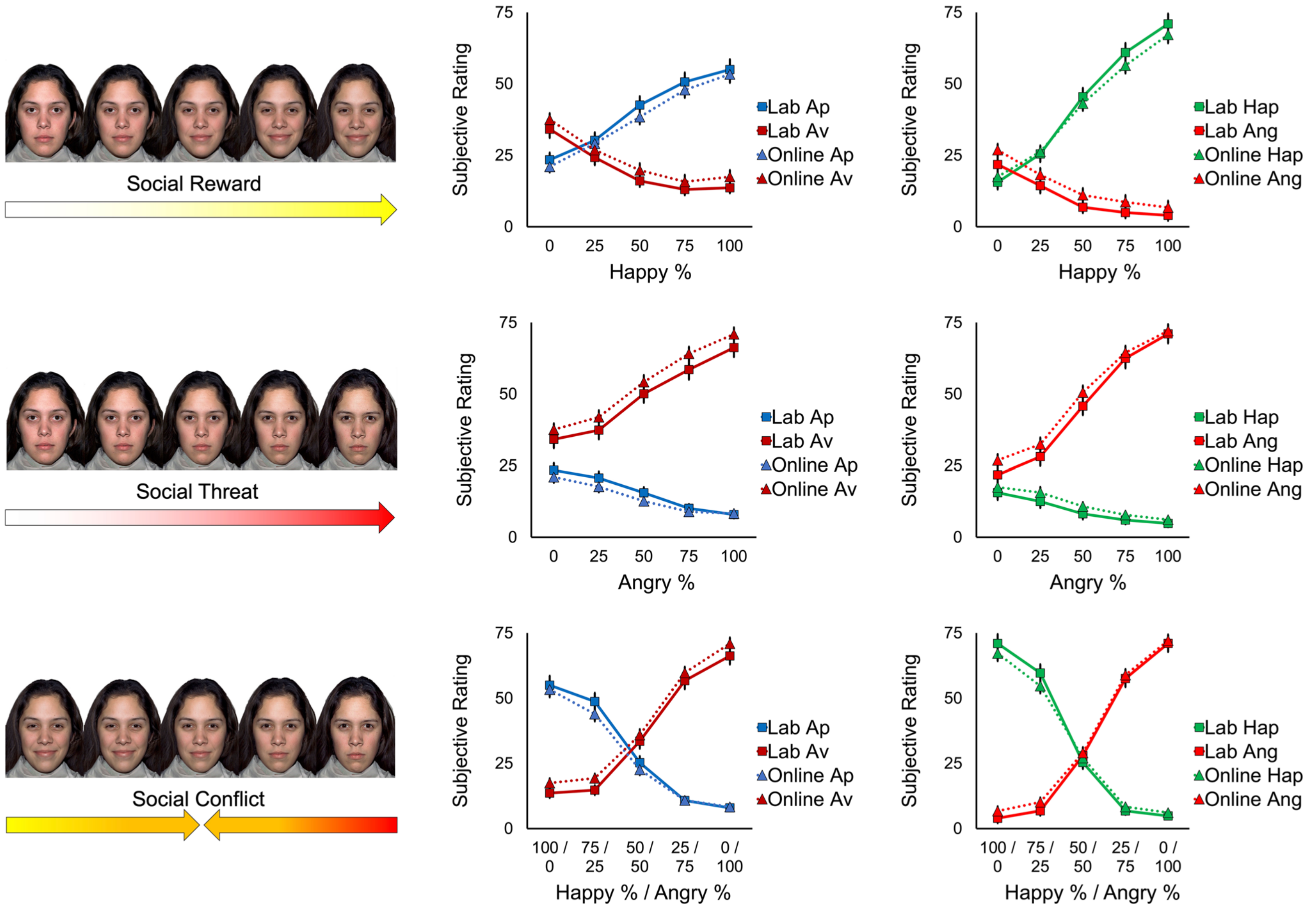

Fig. 2 Replicating and Comparing Task Effects in the Social Approach-Avoidance Paradigm *Legend:* Two independent samples of participants completed the same social approach-avoidance paradigm (SAAP) administered either in the laboratory (solid lines and square markers) or online (dotted lines and triangle markers). As described in the main text, we observed robust SAAP task effects (all $ps < 0.001$), which did not significantly differ between the laboratory sample and online sample (all $ps > 0.22$). Note: *** $p < 0.001$. Ap = Approach, Av = Avoid, Hap = Happiness, Ang = Anger

Internal consistency of SAAP sensitivity measures

For approach-avoidance motivational sensitivity, we generally observed overall strong internal consistency in both samples, which required relatively few slope estimates to reach acceptable levels (see Figure S1). For approach motivational sensitivity, internal consistency estimates ranged from acceptable to excellent for the social reward model (Lab: ω = 0.82, Online: ω = 0.74), social threat model (Lab: ω = 0.83, Online: ω = 0.81), and social reward-threat conflict model (Lab: ω = 0.92, Online: ω = 0.88). For avoidance motivational sensitivity, internal consistency estimates ranged from questionable to excellent for the social reward model (Lab: ω = 0.73, Online: ω = 0.58), social threat model (Lab: ω = 0.85, Online: ω = 0.69), and social reward-threat conflict model (Lab: ω = 0.90, Online: ω = 0.82). For the difference in approach motivational sensitivity and avoidance motivational sensitivity (i.e., Rating × Morph interaction term), internal consistency estimates ranged from acceptable to excellent for the social reward model (Lab: ω = 0.80, Online: ω = 0.70), social threat model (Lab: ω = 0.86, Online: ω = 0.72), and social reward-threat conflict model (Lab: ω = 0.93, Online: ω = 0.88).

For emotion perceptual sensitivity, we also generally observed strong internal consistency estimates in the lab sample and online sample, which required relatively few slope estimates to reach acceptable levels (see Figure S1). For happiness perceptual sensitivity, internal consistency estimates ranged from acceptable to good for the social reward model (Lab: ω = 0.87, Online: ω = 0.76), social threat model (Lab: ω = 0.83, Online: ω = 0.86), and social reward-threat conflict model (Lab: ω = 0.87, Online: ω = 0.83). For anger perceptual sensitivity, internal consistency estimates ranged from questionable to good for the social reward model (Lab: ω = 0.82, Online: ω = 0.76), social threat model (Lab: ω = 0.80, Online: ω = 0.68), and social reward-threat conflict model (Lab: ω = 0.82, Online: ω = 0.76). For the difference in happiness perceptual sensitivity and anger perceptual sensitivity (i.e., Rating × Morph interaction), internal consistency estimates ranged from questionable to good for the social reward model (Lab: $\omega = 0.83$, Online: $\omega = 0.76$), social threat model (Lab: $\omega = 0.78$, Online: $\omega = 0.65$), and social reward-threat conflict model (Lab: $\omega = 0.88$, Online: $\omega = 0.85$).

On average, approach-avoidance motivational and emotional perceptual sensitivity metrics computed using adjacent contrast scores also exhibited similarly strong levels of internal consistency (Lab: $\omega = 0.82$, Online: $\omega = 0.72$). Similarly, ratings in response to each morphed facial expression exhibited internal consistency estimates ranging from acceptable to good on average (Lab: $\omega = 0.84$; Online: $\omega = 0.78$). Importantly, ratings for unmorphed stimuli (e.g., $100\%_{\text{Happy}}$) exhibited highly similar internal consistency estimates (Lab: $\omega = 0.86$; Online: $\omega = 0.79$) compared to ratings for morphed stimuli on average (e.g., $50\%_{\text{Happy}}$; Lab: $\omega = 0.83$; Online: $\omega = 0.79$).

Test–retest reliability of SAAP sensitivity measures

Prior to conducting test–retest analyses, we first tested if participants who completed both online sessions ($n = 31$) exhibited statistically significant differences in SAAP sensitivity measures compared to participants who only completed the first online session ($n = 29$). We did not observe statistically significant differences between these two groups of participants in either approach-avoidance motivational sensitivity (all $ps > 0.57$) or emotion perceptual sensitivity (all $ps > 0.83$).

For approach-avoidance motivational sensitivity, we observed uniformly strong test–retest reliability (see Figure S2). Approach motivational sensitivity and avoidance motivational sensitivity both exhibited acceptable/good test–retest reliability across the social reward model, social threat model, and social reward-threat conflict model (all ICCs > 0.74; see Supplemental Information for details). Similarly, the difference between approach and avoidance motivational sensitivity (i.e., the Rating × Morph interaction), also exhibited good test–retest reliability across the social reward model, social threat model, and social reward-threat conflict model (all ICCs > 0.83; for detailed results, see Supplemental Information).

For emotion perceptual sensitivity, we observed uniformly strong test–retest reliability (see Figure S2). Happiness perceptual sensitivity and anger perceptual sensitivity both exhibited good test–retest reliability across the social reward model, social threat model, and social reward-threat conflict model (all ICCs > 0.77; see Supplemental Information for details). Similarly, the difference between happiness and anger perceptual sensitivity (i.e., the Rating × Morph interaction) also exhibited good test–retest reliability across the social reward model, social threat model, and social reward-threat conflict model (all ICCs > 0.80; for detailed results, see Supplemental Information).

On average, approach-avoidance motivational and emotional perceptual sensitivity metrics computed using adjacent contrast scores also exhibited similarly strong test–retest reliability (ICC = 0.81). Similarly, ratings in response to each morphed facial expression exhibited strong test–retest reliability on average (ICC = 0.81). Moreover, unmorphed stimuli (e.g., $100\%_{\text{Happy}}$) exhibited similar test–retest reliability (ICC = 0.80) compared to morphed stimuli on average (e.g., $50\%_{\text{Happy}}$; ICC = 0.81).

Individual differences in SAAP sensitivity measures

For both the lab and online samples, we observed substantial individual differences in the both the *magnitude* of motivational and perceptual ratings (i.e., random intercepts) and the *sensitivity* of motivational and perceptual ratings (i.e., random slopes; see Fig. 3). Across all models, both the inclusion of random intercepts (relative to no random effects) and the inclusion of random slopes (relative to random intercepts only) robustly improved model fit indices (all $\chi^2 > 70.46$, all $ps < 0.001$; see Supplemental Information for details). Notably, individual differences in SAAP sensitivity measures were generally normally distributed with minor exceptions (see Figure S3).

Discussion

Overall, Study 1 demonstrated that the SAAP reliably measures sensitivity to linear changes in social reward and/or social threat signal intensity. Moreover, using multiple computation methods, sensitivity measures in the SAAP exhibited sufficiently strong psychometric properties to justify future individual differences research in larger samples. To briefly summarize, approach-avoidance motivational sensitivity and emotion perceptual sensitivity in the SAAP: 1) exhibited robust effects that independently replicated across samples, 2) did not exhibit statistically significant differences between laboratory and online administration, 3) generally exhibited strong internal consistency, 4) uniformly exhibited strong test–retest reliability, and 5) exhibited substantial individual differences. Taken together, these results demonstrate that the SAAP exhibits the necessary psychometric properties for large-scale individual differences research that can assess the overall quality of the SAAP as an assessment tool.

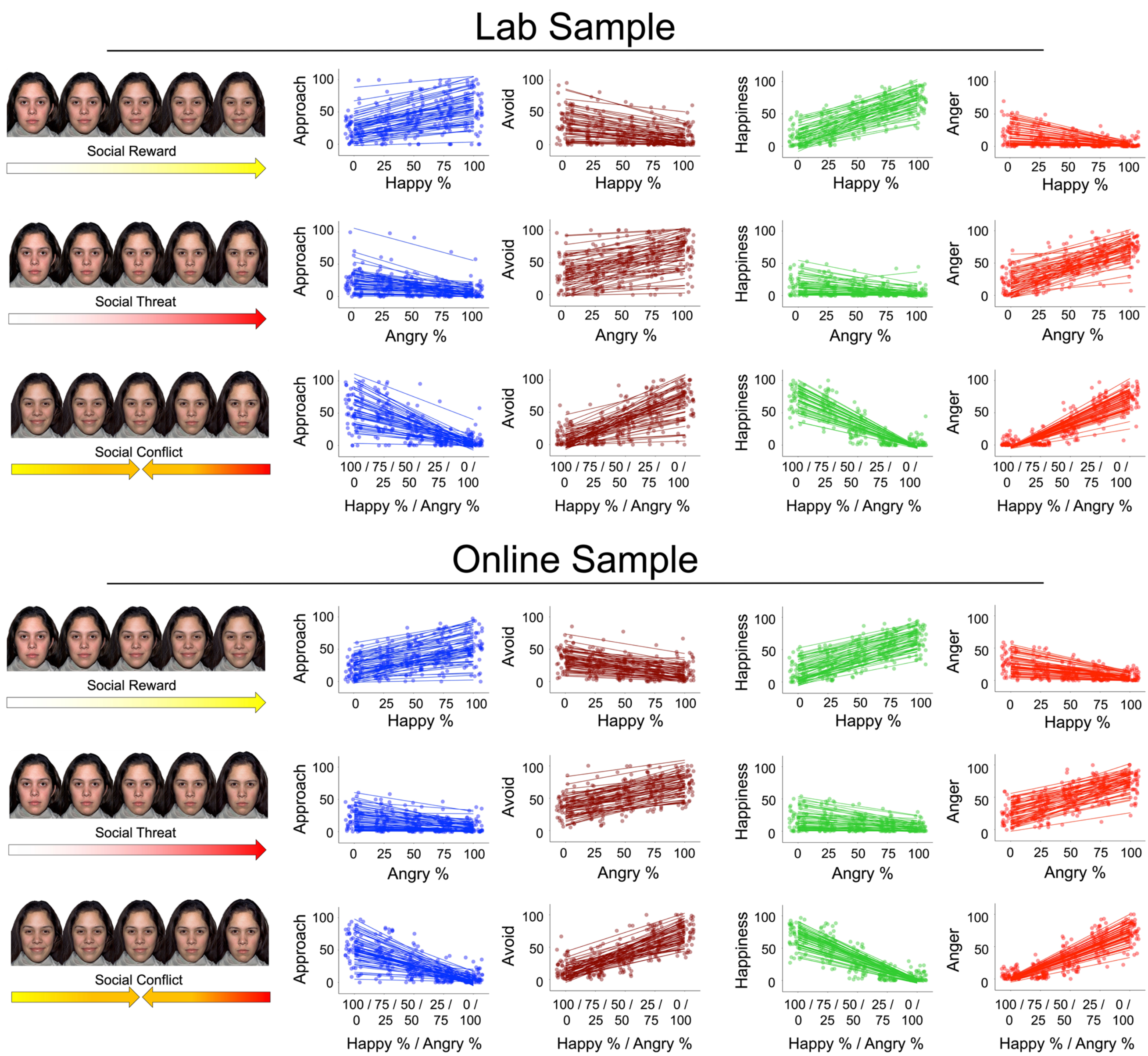

Fig. 3 Individual Differences in Sensitivity Scores.*Legend:* Within the lab sample and online sample, random slope plots illustrate between-subject variability in approach motivational sensitivity (blue lines), avoidance motivational sensitivity (dark red lines), happiness perceptual sensitivity (green lines), and anger perceptual sensitivity (light red lines). Note: Lines represent the linear slope of best fit for individual participants and dots represent average ratings of each morphed facial expression for individual participants

Study 2

Introduction

Overall, Study 1 established that morphed facial expressions in the SAAP reliably modulate approach-avoidance motivation and emotion perception, which are *directly* influenced by varying degrees of happiness and/or anger. However, it remains unclear if the SAAP can also be used to measure the sensitivity of more complex interpersonal judgements that are more *indirectly* influenced by varying degrees of happiness and/or anger. Specifically, previous research demonstrates that subtle signals of happiness and/or anger within facial expressions are also used to make more elaborative interpersonal judgements about other individuals (Todorov, 2008). For example, subtle signals of happiness and/or anger are used to generate inferences about another individual's trustworthiness and/or potential for aggression, which influences a variety of social outcomes and social behaviors (Todorov et al., 2015). Therefore, it is important to determine if the SAAP generalizes to other social judgements that are more *indirectly* sensitive to subtle signals of happiness and/or anger. To address this issue in Study 2, we used the SAAP to characterize and compare the sensitivity of multiple social judgements in addition to basic motivation and emotion perception.

Based on the emotion overgeneralization hypothesis of face processing, interpersonal inferences about other individuals are based on the degree to which structural facial features resemble expressions of happiness and/or anger (Knutson, 1996; Montepare & Dobish, 2003; Todorov, 2011; Zebrowitz, 2018). Importantly, these more complex social judgments are reliably generated in response to affectively *neutral* faces, which do not contain overt emotional expression information (Todorov et al., 2015). For example, the degree to which a person is evaluated as trustworthy is strongly correlated with the degree to which that person's *neutral* facial expression is perceived as happy (Todorov, 2008). Similarly, evaluations of aggressiveness are robustly correlated with the degree to which that person's *neutral* facial expression is perceived as angry (Todorov, 2008). Moreover, even seemingly invariant evaluations of the same person such as physical attractiveness are surprisingly sensitive to subtle differences in facial expressions from the same individual (Jenkins et al., 2011). Thus, more complex interpersonal judgements can be rapidly generated by relying on subtle signals of happiness and/or anger, albeit in a manner that is often inaccurate in regards to the individual's personality traits (Todorov et al., 2015).

Although interpersonal judgements are sensitive to subtle signals of happiness and/or anger, there are nevertheless comparative differences in sensitivity among these more complex dimensions of face evaluation. For example, facial expressions that are perceived as happy are also evaluated as more trustworthy, and to a lesser extent, more attractive (Todorov, 2011). Similarly, facial expressions perceived as angry are evaluated as more aggressive, and to a lesser extent, more dominant (Todorov, 2011). Notably, evaluations of attractiveness and aggressiveness are non-selectively associated with both signals of happiness and/or anger as well as characteristics such as facial symmetry and/or maturity (Todorov et al., 2015). Despite these comparative differences in sensitivity, however, individual differences in interpersonal judgements tend to be intercorrelated across face evaluation dimensions, which suggests some overlapping influence by subtle social signals (Todorov, 2011).

Based on this literature, we aimed to replicate and extend the results of Study 1 by assessing if the SAAP generalized to more complex interpersonal judgements that are more indirectly influenced by subtle signals of happiness and/or anger. In the SAAP, we hypothesized that these more complex interpersonal judgements would also exhibit sensitivity to varying interpersonal reward, social threat, and social reward-threat conflict. Given the more indirect influence by social signals, we also hypothesized that more complex interpersonal judgements would exhibit comparatively weaker sensitivity than approach-avoidance motivation and emotion perception ratings. Additionally, based on previous research, we hypothesized that ratings of physical attractiveness and dominance would exhibit the lowest degree of sensitivity among these more complex interpersonal judgements.

Transparency and openness

Study 2 was not preregistered prior to data collection and analysis. De-identified data and all analysis syntax for Study 2 is available in an open-access data repository (https://osf.io/yh42j/ bb2c5bf777144b39a75c8a6ed158007d). Morphed facial expression stimuli can be made freely available to researchers who receive written permission to access the NimStim set (Tottenham et al., 2009).

We report how we determined our sample size, all data exclusions, and all manipulations in Study 2. All participants provided written informed consent and study procedures were conducted in accordance with local IRB guidelines for Harvard Medical School (Protocol: #16–0624). All study procedures complied with the Helsinki Declaration as revised in 2008.

Methods

Participants

Sample size for Study 2 was determined based on a pilot study of 67 undergraduate students from the University of Miami. In this pilot study, the smallest multivariate difference in sensitivity among face evaluation dimensions exhibited a medium effect size ($f=0.25$). Based on this effect size, a sample size of 84 participants would provide 95% power to detect multivariate differences in sensitivity across face evaluation dimensions.

Based on these power analyses, we recruited an online sample of 93 healthy adults via Prolific Academic from the Boston community (74.20% Female; Age: $M=41.39$ years, $SD=11.86$ years; 81.70% Caucasian, 3.20% Black, 5.40% Asian, 8.60% Multi-Racial/Other, 1.10% Did Not Report). Participants were recruited as part of a healthy control sample within larger research protocol on Developmental Prosopagnosia. Thus, participants in Study 2 denied any current or history of psychiatric disorders and successfully completed all effort checks throughout the online testing session. Additionally, all participants were required to exhibit intact face recognition as defined by performance greater than 70% accuracy on the famous faces task (Duchaine & Nakayama, 2006) and a score less than 65 on the Prosopagnosia Index questionnaire (Shah et al., 2015). Participants were provided with monetary payment ($10 per hour) in exchange for participation.

Stimuli and task

Social Approach-Avoidance Paradigm

In Study 2, we used the same morphed facial expressions as Study 1, which participants evaluated using a 10-point dynamic visual scale (0 = *Not at all*; 10 = *Extremely*). In addition to approach-avoidance motivation and perceived happiness/anger, participants also rated facial expressions based on trustworthiness, physical attractiveness, aggressiveness, and dominance. To minimize fatigue, participants completed all eight rating dimensions for two male faces and two female faces, rather than four male faces and four female faces as in Study 1. To minimize task switching and order effects, rating evaluations were blocked and the order of blocks was randomized across participants. In total, participants completed 384 ratings (i.e., 48 approach, 48 avoid, 48 happiness, 48 anger, 48 trustworthiness, 48 attractiveness, 48 aggressiveness, and 48 dominance).

Data analytic strategy

All analyses were conducted using SPSS software ver. 24.0 (IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM).

Characterizing the sensitivity of face evaluation dimensions

First, we aimed to confirm if each evaluation dimension was sensitive to parametrically varying social signals. Specifically, we tested if subjective ratings for each evaluation dimension exhibited linear modulation as a function of social reward, social threat, or social reward-threat conflict. To this end, we used RM-ANOVA models with linear polynomial contrasts to separately test if subjective ratings for each dimension exhibited reliable linear modulation as a function of social signals. Thus, significant sensitivity in the SAAP was defined as a linear slope value that significantly differed from 0. Additionally, on an exploratory basis, we characterized the intercorrelation of sensitivity measures across face evaluation dimensions to determine the degree to which social signals exert overlapping influences on these processes (see Supplemental Information).

Comparing the sensitivity of face evaluation dimensions

Second, we directly compared the sensitivity of face evaluation dimensions as a function of varying social signals. To this end, we submitted subjective ratings to an 8 (<u>Dimension</u>: [Approach, Avoidance, Happiness, Anger, Trustworthiness, Attractiveness, Aggression, and Dominance]) × Morph (Linear: -2, -1, 0, 1, 2) RM-ANOVA model. To decompose significant omnibus interactions, we separately compared the magnitude of sensitivity to social signals for positive valence evaluations (i.e., approach, happiness, trustworthiness, and attractiveness) and negative valence evaluations (i.e., avoidance, anger, aggressiveness, and dominance). Within these positive valence models and negative valence models, we conducted post-hoc pairwise comparisons between face evaluation dimensions with Bonferroni correction to account for multiple comparisons.

Results

Characterizing the sensitivity of face evaluation dimensions

We observed significant task-related effects across all dimensions of face evaluation. Specifically, all face evaluation dimensions exhibited significant sensitivity within the social reward model, social threat model, and social reward-threat conflict model with large effect sizes (all ps < 0.001, all $\eta_p^2 > 0.36$; see Table S1). Overall, sensitivity to social signals generally exhibited a moderate degree of intercorrelation across face evaluation dimensions, which varied across evaluation dimensions and models (see Supplemental Information).

Comparing the sensitivity of face evaluation dimensions

Social reward model

Within the omnibus model testing an 8 (Dimension) × Morph (Linear) interaction, face evaluation dimensions significantly differed in sensitivity to varying degrees of social reward ($F_{(7, 89)} = 282.91$, $p < 0.001$, $\eta_p^2 = 0.75$; see Figure S4). Follow-up analyses demonstrated that both positive valence dimensions ($F_{(3, 93)} = 171.47$, $p < 0.001$, $\eta_p^2 = 0.64$; see Fig. 4) and negative valence dimensions ($F_{(3, 93)} = 6.85$, $p < 0.001$, $\eta_p^2 = 0.07$; see Fig. 4) significantly differed in sensitivity to varying degrees of social reward.

For *positive* valence dimensions, post-hoc comparisons revealed significant pairwise differences in sensitivity (i.e., happiness > approach > trustworthiness > attractiveness; see Fig. 4). As social reward increased, happiness ratings *increased* to a greater degree compared to approach, trustworthiness, and attractiveness ratings (all Fs > 17.51, all ps < 0.001, corrected). Additionally, approach ratings *increased* to a greater degree compared to both trustworthiness and attractiveness ratings (both Fs > 141.15, both ps < 0.001, corrected). Finally, trustworthiness ratings

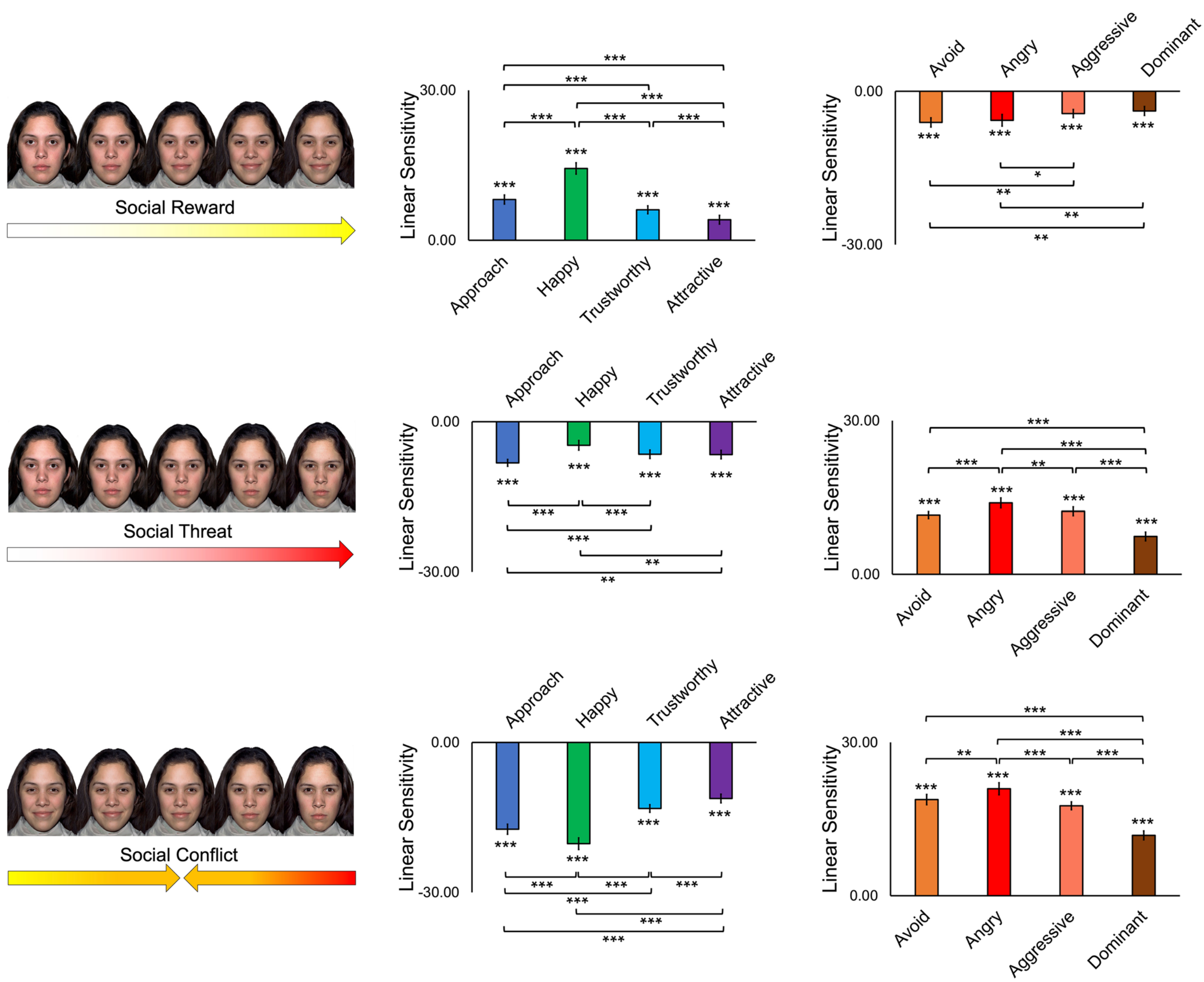

Fig. 4 Generalization and Comparison of Social Sensitivity across Face Evaluation Dimensions. *Legend:* All face evaluation dimensions exhibited significant sensitivity to varying social signals in the social approach-avoidance paradigm (SAAP; all $ps < 0.001$). Additionally, face evaluation dimensions significantly differed in the degree of sensitivity (all $ps < 0.001$). Post-hoc pairwise comparisons with Bonferroni correction were used to compare sensitivity between positive valence dimensions (left column) and negative valence dimensions (right column). Note: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, Bonferroni corrected

increased to a greater degree than attractiveness ratings ($F(_{1, 95}) = 43.64$, $p < 0.001$, corrected).

For *negative* valence dimensions, post-hoc comparisons revealed significant pairwise differences in sensitivity (i.e., anger = avoidance > aggressiveness = dominance; see Fig. 4). As social reward increased, anger ratings *decreased* to a greater degree compared to both aggressiveness and dominance ratings (both $Fs > 11.54$, both $ps < 0.006$, corrected), but did not significant differ in sensitivity relative to avoidance ratings ($F(_{1, 95}) = 0.32$, $p = 1.00$, corrected). Additionally, avoidance ratings *decreased* to a greater degree compared to dominance ratings ($F(_{1, 95}) = 8.69$, $p = 0.02$, corrected), whereas the comparison to aggressiveness ratings did not survive multiple comparison correction ($F(_{1, 95}) = 6.80$, $p = 0.06$, corrected).

Social threat model

Within the omnibus model testing an 8 (Dimension) × Morph (Linear) interaction, face evaluation dimensions significantly differed in sensitivity to varying degrees of social threat ($F(_{7, 89}) = 437.35$, $p < 0.001$, $\eta_p^2 = 0.82$; see Figure S4). Follow-up analyses demonstrated that both positive valence dimensions ($F(_{3, 93}) = 25.12$, $p < 0.001$, $\eta_p^2 = 0.21$; see Fig. 4) and negative valence dimensions ($F(_{3, 93}) = 41.96$, $p < 0.001$, $\eta_p^2 = 0.31$; see Fig. 4) significantly differed in sensitivity to varying degrees of social threat.

For *positive* valence dimensions, post-hoc comparisons revealed significant pairwise differences in sensitivity (i.e., approach > trustworthiness = attractiveness > happiness;

see Fig. 4). As social threat increased, approach ratings *decreased* to a greater extent relative to happiness, trustworthiness, and attractiveness ratings (all Fs > 25.60, all ps < 0.001, corrected). Additionally, trustworthiness and attractiveness ratings both *decreased* to a greater extent relative to happiness ratings (both Fs > 8.85, both ps < 0.02, corrected). However, trustworthiness did not significantly differ in sensitivity compared to attractiveness ratings ($F_{(1, 95)} = 0.17$, $p = 1.00$, corrected).

For *negative* valence dimensions, post-hoc comparisons revealed significant pairwise differences in sensitivity (i.e., anger > avoidance > aggressiveness = dominance; see Fig. 4). As social threat increased, anger ratings *increased* to a greater extent relative to avoidance, aggressiveness, and dominance ratings (all Fs > 11.80, all ps < 0.005, corrected). Additionally, avoidance and aggressiveness ratings both *increased* to a greater extent relative to dominance ratings (both Fs > 32.67, both ps < 0.001, corrected). In contrast, avoidance ratings did not significantly differ in sensitivity compared to aggressiveness ratings $F_{(1, 95)} = 0.41$, $p = 1.00$, corrected).

Social reward-threat conflict model

Within the omnibus model testing an 8 (Dimension) × Morph (Linear) interaction, face evaluation dimensions significantly differed in sensitivity to varying degrees of co-occurring social reward and social threat ($F_{(7, 89)} = 746.07$, $p < 0.001$, $\eta_p^2 = 0.96$; see Fig. 4G). Follow-up analyses demonstrated that both positive valence dimensions ($F_{(3, 93)} = 127.68$, $p < 0.001$, $\eta_p^2 = 0.57$; see Fig. 4H) and negative valence dimensions ($F_{(3, 93)} = 62.32$, $p < 0.001$, $\eta_p^2 = 0.40$; see Figure S4) significantly differed in sensitivity to varying degrees of co-occurring social reward and social threat.

For *positive* valence dimensions, post-hoc comparisons revealed significant pairwise differences in sensitivity (i.e., happiness > approach > trustworthiness > attractiveness; see Fig. 4). As social reward decreased and co-occurring social threat increased, happiness ratings *decreased* to a greater extent relative to approach, trustworthiness, and attractiveness ratings (all Fs > 28.59, all ps < 0.001, corrected). Additionally, approach ratings *decreased* to a greater extent compared to both trustworthiness and attractiveness ratings (both Fs > 65.56, both ps < 0.001, corrected). Finally, trustworthiness ratings *decreased* to a greater extent compared to attractiveness ratings ($F_{(1, 95)} = 26.86$, $p < 0.001$, corrected).

For *negative* valence dimensions, post-hoc comparisons revealed significant pairwise differences in sensitivity (i.e., anger > avoidance > aggressiveness > dominance; see Fig. 4). As social reward decreased relative to co-occurring social threat, anger ratings *increased* to a greater extent relative to avoidance, aggressiveness, and dominance ratings (all Fs > 7.39, all ps < 0.05, corrected). Additionally, avoidance and aggressiveness ratings both *increased* to a greater extent relative to dominance ratings (both Fs > 103.42, both ps < 0.001, corrected). However, avoidance ratings did not significantly differ in sensitivity compared to aggressiveness ratings ($F_{(1, 95)} = 3.62$, $p = 0.36$, corrected).

Discussion

In Study 2, we replicated and extended Study 1 by generalizing the SAAP to more complex interpersonal judgments that are *indirectly* modulated by signals of happiness and/or anger. Consistent with our hypotheses, more complex interpersonal judgements (trustworthiness, physical attractiveness, aggressiveness, and dominance) were each sensitive to parametric changes in social signals. Also consistent with our hypotheses, these more complex emotional judgements were generally less sensitive to varying social signals than approach-avoidance motivation and emotion perception. Finally, sensitivity to social signals generally exhibited a moderate degree of intercorrelation across face evaluation dimensions, which suggests that these dimensions capture both overlapping and unique influences by social signals. Taken together, these results demonstrate that in addition to basic motivation and emotion perception processes, more complex interpersonal judgements are also sensitive to linear changes in social reward and/or social threat intensity, albeit to varying degrees.

General discussion

Across two studies and three independent samples, we demonstrated that approach-avoidance motivation, emotion perception, and more complex interpersonal judgements are sensitive to linear changes in social signal intensity. Study 1 demonstrated that sensitivity to linear changes in social signal intensity exhibited strong psychometric properties as evidenced by 1) robust effect sizes and replication *across* samples, 2) strong internal consistency and test–retest reliability *within* participants, and 3) robust individual differences *between* participants. Study 2 demonstrated that more complex interpersonal judgements, which are more indirectly influenced by subtle signals of happiness and/or anger, are also sensitive to linear changes in social signal intensity to varying degrees. Together, these studies demonstrate that the SAAP reliably assays the sensitivity of multiple social processes with the necessary psychometric properties to justify conducting more large-scale individual differences research on the quality of the SAAP as an assessment tool for better understanding social behavior.

Given the replication crisis in psychology, it is imperative to develop experimental paradigms that produce replicable measures with strong psychometric properties to facilitate more rigorous mechanistic and individual differences research (Parsons et al., 2019). Addressing this issue, we demonstrated that SAAP sensitivity measures are internally consistent and temporally stable within participants, while also exhibiting sufficient between-participant variability required to detect associations with other measures. Moreover, the strong temporal stability of SAAP sensitivity measures provides clinical researchers with the ability to reliably measure intervention-related changes in sensitivity to varying social signals. Finally, although we did not observe statistically significant differences in SAAP task effects between laboratory and online task administration, more trials may be required in online protocols to ensure uniformly strong internal consistency across measures. Together, these results suggest that the SAAP can be used to reliably assay motivational and emotional sensitivity to linear changes in social signal intensity with the necessary psychometric properties to justify conducting individual differences research at a larger scale.

However, it is also important to acknowledge that these results do not provide a comprehensive examination of the overall quality of the SAAP as a measure of social cognition. To address this issue, additional work will be required in much larger samples that directly test the convergent validity and discriminant validity of SAAP task measures. For example, it will be important for future research to assess the degree to which SAAP task measures exhibit meaningful associations with other, more established measures of social cognition (e.g., Reading the Mind in the Eyes; Baron-Cohen et al., 2001). Similarly, additional measures and methodologies will be necessary to test discriminant validity using multitrait-multimethod (MTMM) matrices in line with current assessment standards (Rönkkö & Cho, 2022). Finally, to fully validate the SAAP as a social cognition measure, it will be necessary to establish norm values in much larger, representative samples. Therefore, the current study only provides a necessary, but not sufficient, first step towards empirically validating the SAAP as a rigorous measure of social cognition.

In addition to approach-avoidance motivation and emotion perception, Study 2 demonstrated that SAAP can be used to measure the sensitivity of more complex interpersonal judgements that also contribute to social behavior. Specifically, interpersonal judgements of trustworthiness, attractiveness, aggressiveness, and dominance all exhibited robust sensitivity to varying social signals in the SAAP. Notably, we also observed comparative differences in sensitivity among face evaluation dimensions. Specifically, approach-avoidance motivation and emotion perception generally exhibited the greatest degree of sensitivity to linear changes in social signal intensity, trustworthiness and aggressiveness generally exhibited an intermediate degree of sensitivity, and physical attractiveness and dominance exhibited the smallest degree of sensitivity. These comparative differences in sensitivity to subtle signals of happiness and/or anger align with previous research using neutral facial expressions (Todorov, 2015). Specifically, previous studies demonstrate that evaluations of trustworthiness and aggressiveness are predominantly influenced by affective signals such as happiness and/or anger, whereas evaluations of physical attractiveness and dominance are jointly influenced by both affective signals and non-affective signals such as facial symmetry and facial maturity (DeBruine, 2005; Golle et al., 2014; Morrison et al., 2013). Thus, Study 2 demonstrates that these more complex interpersonal judgements are also sensitive to linear changes in social signal intensity in the SAAP, albeit to varying degrees.

Despite these important strengths, it is also important to highlight several limitations of the current studies. First, to minimize the confounding influence of racial identity on face evaluation (Paulus & Wentura, 2014), we exclusively utilized Caucasian faces in Study 1 and Study 2. To address this limitation, we recently developed morphed facial expressions using a more racially representative set of faces, which we are currently working to validate with the SAAP. Second, given the extremely large number of potential face judgements (e.g., competence, intelligence, etc.), it was not possible to exhaustively assess if the SAAP generalizes to all possible face judgements. Therefore, we recommended that researchers first validate that the SAAP generalizes to untested face judgements prior characterizing mechanisms and/or individual differences in these face judgements. Third, there is a smaller body of empirical research supporting the ecological validity of using linearly interpolated morphs to approximate facial expressions commonly observed in daily life. On one hand, previous research has established that linearly interpolated morphed facial expressions adequately capture course changes in facial action units (AUs) underlying emotional facial expressions, which directly corresponds to perceived emotional intensity of linearly interpolated morphs rated by human observers (Calvo et al., 2018). Additionally, linearly interpolated morphs used in the current study (Neutral-Happy, Neutral-Angry, and Happy-Angry) exhibit similar changes in facial landmarks that do not differ compared to dynamic spontaneous expressions of emotion (Korolkova, 2018). On the other hand, research also demonstrates that linear interpolation of facial expressions does not fully capture more complex, non-linear spatiotemporal shifts in facial action units that dynamically unfold during spontaneous emotion expression (Cosker et al., 2010; Dobs et al., 2014; Krumhuber et al., 2023). Therefore, it will be important for future research to examine the sensitivity of motivation and emotion perception in response to non-linear changes in social signal intensity.

Despite these limitations, however, the replication crisis in psychology has underscored the importance of developing experimental paradigms that produce replicable measures with strong psychometric properties. Taken together, these findings demonstrate that the sensitivity of motivational and social processes to linear changes in social signal intensity can be rigorously and reliably assayed with the SAAP. Although future research with larger samples is needed to establish convergent validity, discriminant validity, and norm values of the SAAP as a measure of social cognition, the results of the current study suggest that measuring the sensitivity of motivational and emotional processes may provide insights into social behavior.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11031-024-10076-z.

Funding This research was supported by a VA RR&D SPiRE grant (1I21RX003727-01A1) to ME, an R01 grant from the National Eye Institute to JD (R01EY032510-02), and an Office of Academic Affiliations fellowship to TCE.

References

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the Enigmatic Face: The Importance of Facial Dynamics in Interpreting Subtle Facial Expressions. *Psychological Science, 16*(5), 403–410. https://doi.org/10.1111/j.0956-7976.2005.01548.x

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(2), 241–251.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest, 20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beevers, C. G., Wells, T. T., Ellis, A. J., & Fischer, K. (2009). Identification of emotionally ambiguous interpersonal stimuli among dysphoric and nondysphoric individuals. *Cognitive Therapy and Research, 33*(3), 283–290. https://doi.org/10.1007/s10608-008-9198-6

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H., & White, J. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution, 24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Calvo, M. G., Fernández-Martín, A., Recio, G., & Lundqvist, D. (2018). Human observers and automated assessment of dynamic emotional facial expressions: KDEF-dyn database validation. *Frontiers in Psychology, 9*, 397727.

Calvo, M. G., Gutiérrez-García, A., Fernández-Martín, A., & Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *Journal of Nonverbal Behavior, 38*, 549–567.

Carroll, J. M., & Russell, J. A. (1997). Facial expressions in Hollywood's protrayal of emotion. *Journal of Personality and Social Psychology, 72*(1), 164.

Chen, C., & Jack, R. E. (2017). 2017/10/01/). Discovering cultural differences (and similarities) in facial expressions of emotion. *Current Opinion in Psychology, 17*, 61–66. https://doi.org/10.1016/j.copsyc.2017.06.010

Cho, E., & Kim, S. (2015). Cronbach's Coefficient Alpha: Well Known but Poorly Understood. *Organizational Research Methods, 18*(2), 207–230. https://doi.org/10.1177/1094428114555994

Cooper, S. R., Gonthier, C., Barch, D. M., & Braver, T. S. (2017). The role of psychometrics in individual differences research in cognition: A case study of the AX-CPT. *Frontiers in psychology, 8*, 290877

Cosker, D., Krumhuber, E., & Hilton, A. (2010, July). Perception of linear and nonlinear motion properties using a FACS validated 3D facial model. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (pp. 101–108).

DeBruine, L. M. (2005). Trustworthy but not lust-worthy: context-specific effects of facial resemblance. *Proceedings of the Biological Sciences, 272*(1566), 919–922. https://doi.org/10.1098/rspb.2004.3003

Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in behaviour: mixed-effect modelling approaches [https://doi.org/10.1111/1365-2656.12013]. *Journal of Animal Ecology, 82*(1), 39–54. https://doi.org/10.1111/1365-2656.12013.

Dobs, K., Bülthoff, I., Breidt, M., Vuong, Q. C., Curio, C., & Schultz, J. (2014). Quantifying human sensitivity to spatio-temporal information in dynamic faces. *Vision Research, 100*, 78–87.

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

Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences, 111*(15), E1454–E1462.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*(4), 576–585. https://doi.org/10.1016/j.neuropsychologia.2005.07.001

Evans, T. C., & Britton, J. C. (2018). Sep). Improving the psychometric properties of dot-probe attention measures using response-based computation. *Journal of Behavior Therapy and Experimental Psychiatry, 60*, 95–103. https://doi.org/10.1016/j.jbtep.2018.01.009

Evans, T. C., & Britton, J. C. (2020). Social avoidance behaviour modulates automatic avoidance actions to social reward-threat conflict. *Cognition and Emotion, 34*(8), 1711–1720. https://doi.org/10.1080/02699931.2020.1787353

Evans, T. C., Esterman, M., & Britton, J. C. (2022). Social avoidance behavior modulates motivational responses to social reward-threat conflict signals: A preliminary fMRI study. *Cognitive, Affective, & Behavioral Neuroscience*. https://doi.org/10.3758/s13415-022-01031-x

Evans, T. C., Walukevich, K. A., Seager, I., & Britton, J. C. (2018). A psychometric comparison of anxiety-relevant attention measures. *Anxiety Stress Coping, 31*(5), 539–554. https://doi.org/10.1080/10615806.2018.1489536

Frith, C. (2009). Role of facial expressions in social interactions. *Philos Trans R Soc Lond B Biol Sci, 364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Fröhner, J. H., Teckentrup, V., Smolka, M. N., & Kroemer, N. B. (2019). Jul 15). Addressing the reliability fallacy in fMRI: Similar group effects may arise from unreliable individual effects. *NeuroImage, 195*, 174–189. https://doi.org/10.1016/j.neuroimage.2019.03.053

Golle, J., Mast, F. W., & Lobmaier, J. S. (2014). Something to smile about: The interrelationship between attractiveness and emotional expression. *Cognition and Emotion, 28*(2), 298–310. https://doi.org/10.1080/02699931.2013.817383

Goodhew, S. C., & Edwards, M. (2019). 2019/03/01/). Translating experimental paradigms into individual-differences research: Contributions, challenges, and practical recommendations. *Consciousness and Cognition, 69*, 14–25. https://doi.org/10.1016/j.concog.2019.01.008

Goodwin, L. D., & Leech, N. L. (2006). Understanding Correlation: Factors That Affect the Size of r. *The Journal of Experimental Education, 74*(3), 249–266. https://doi.org/10.3200/JEXE.74.3.249-266

Gutiérrez-García, A., & Calvo, M. G. (2014). Social anxiety and interpretation of ambiguous smiles. *Anxiety Stress Coping, 27*(1), 74–89. https://doi.org/10.1080/10615806.2013.794941

Gutiérrez-García, A., & Calvo, M. G. (2016). Oct 30). Social anxiety and perception of (un)trustworthiness in smiling faces. *Psychiatry Research, 244*, 28–36. https://doi.org/10.1016/j.psychres.2016.07.004

Hayden, E. P. (2022). A call for renewed attention to construct validity and measurement in psychopathology research. *Psychological Medicine, 52*(14), 2930–2936. https://doi.org/10.1017/S0033291722003221

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Infantolino, Z. P., Luking, K. R., Sauder, C. L., Curtin, J. J., & Hajcak, G. (2018). Jun). Robust is not necessarily reliable: From within-subjects fMRI contrasts to between-subjects comparisons. *NeuroImage, 173*, 146–152. https://doi.org/10.1016/j.neuroimage.2018.02.024

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition, 121*(3), 313–323. https://doi.org/10.1016/j.cognition.2011.08.001

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in psychology, 1*, 2124.

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior, 20*(3), 165–182.

Korolkova, O. A. (2018). The role of temporal inversion in the perception of realistic and morphed dynamic transitions between facial expressions. *Vision Research, 143*, 42–51.

Krumhuber, E. G., Skora, L. I., Hill, H. C., & Lander, K. (2023). The role of facial movements in emotion recognition. *Nature Reviews Psychology, 2*(5), 283–296.

Kuckertz, J. M., Strege, M. V., & Amir, N. (2017). Intolerance for approach of ambiguity in social anxiety disorder. *Cognition and Emotion, 31*(4), 747–754. https://doi.org/10.1080/02699931.2016.1145105

LaPlante, D., & Ambady, N. (2000). Multiple messages: Facial recognition advantage for compound expressions. *Journal of Nonverbal Behavior, 24*, 211–224.

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), 772.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: An introductory tutorial and free, graphical user interface (LMMgui). *Frontiers in Psychology, 6*, 2. https://doi.org/10.3389/fpsyg.2015.00002

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion, 5*(1), 119–124. https://doi.org/10.1037/1528-3542.5.1.119

Martinez, J. E., Funk, F., & Todorov, A. (2020). Quantifying idiosyncratic and shared contributions to judgment. *Behavior Research Methods, 52*(4), 1428–1444. https://doi.org/10.3758/s13428-019-01323-0

Matsumoto, D., & Hwang, H. C. (2014). Judgments of subtle facial expressions of emotion. *Emotion, 14*(2), 349.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review, 15*(1), 28–50. https://doi.org/10.1177/1088868310366253

Miller, J., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review, 20*(5), 819–858. https://doi.org/10.3758/s13423-013-0404-5

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal behavior, 27*, 237–254.

Morrison, E. R., Morris, P. H., & Bard, K. A. (2013). The stability of facial attractiveness: Is it what you've got or what you do with it? *Journal of Nonverbal Behavior, 37*, 59–67.

Nikitin, J., & Freund, A. M. (2019). The motivational power of the happy face. *Brain Sciences, 9*(1), 6.

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Struhl, M. K., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, Robustness, and Reproducibility in Psychological Science. *Annual Review of Psychology, 73*(1), 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Nunnally, J. C. (1967). Psychometric Theory. New York: McGraw-Hill.

Parsons, S., Kruijt, A.-W., & Fox, E. (2019). Psychological Science Needs a Standard Practice of Reporting the Reliability of Cognitive-Behavioral Measurements. *Advances in Methods and Practices in Psychological Science, 2*(4), 378–395. https://doi.org/10.1177/2515245919879695

Paulus, A., & Wentura, D. (2014). Threatening joy: Approach and avoidance reactions to emotions are influenced by the group membership of the expresser. *Cognition and Emotion, 28*(4), 656–677. https://doi.org/10.1080/02699931.2013.849659

Polit, D. F. (2014). Getting serious about test–retest reliability: a critique of retest research and some recommendations. *Quality of Life Research, 23*(6), 1713–1720. https://doi.org/10.1007/s11136-014-0632-9

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Radke, S., Kalt, T., Wagels, L., & Derntl, B. (2018). Implicit and explicit motivational tendencies to faces varying in trustworthiness and dominance in men. *Frontiers in Behavioral Neuroscience, 12*, 8.

Renard, S. B., de Jong, P. J., & Pijnenborg, G. H. M. (2017). Approach-avoidance of facial affect is moderated by the presence of an observer-irrelevant trigger. *Motivation and Emotion, 41*(2), 265–272. https://doi.org/10.1007/s11031-016-9595-1

Revelle, W., & Zinbarg, R. E. (2009). Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika, 74*(1), 145–154. https://doi.org/10.1007/s11336-008-9102-z

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology, 62*, 23–48. https://doi.org/10.1146/annurev.psych.121208.131647

Roelofs, K., Hagenaars, M. A., & Stins, J. (2010). Facing freeze: Social threat induces bodily freeze in humans. *Psychological Science, 21*(11), 1575–1581. https://doi.org/10.1177/0956797610384746

Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods, 25*(1), 6–14.

Ross, E. D., Gupta, S. S., Adnan, A. M., Holden, T. L., Havlicek, J., & Radhakrishnan, S. (2016). Neurophysiology of spontaneous facial expressions: I. Motor control of the upper and lower face is behaviorally independent in adults. *Cortex, 76*, 28–42.

Scherer, K. R., & Ceschi, G. (2000). Criteria for emotion recognition from verbal and nonverbal expression: Studying baggage loss in the airport. *Personality and Social Psychology Bulletin, 26*(3), 327–339.

Scherer, K. R., & Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion, 7*(1), 113.

Seidel, E.-M., Habel, U., Kirschner, M., Gur, R. C., & Derntl, B. (2010). The impact of facial emotional expressions on behavioral tendencies in women and men. *Journal of Experimental Psychology: Human Perception and Performance, 36*(2), 500–507. https://doi.org/10.1037/a0018169

Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): a self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science, 2*(6), 140343. https://doi.org/10.1098/rsos.140343

Staugaard, S. R. (2010). Threatening faces and social anxiety: a literature review. *Clinical Psychology Review, 30*(6), 669–690. https://doi.org/10.1016/j.cpr.2010.05.001

Stins, J. F., Roelofs, K., Villan, J., Kooijman, K., Hagenaars, M. A., & Beek, P. J. (2011). Walk to me when I smile, step back when I'm angry: emotional faces modulate whole-body approach–avoidance behaviors. *Experimental Brain Research, 212*(4), 603–611. https://doi.org/10.1007/s00221-011-2767-z

Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review, 8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1

Tamir, D. I., & Hughes, B. L. (2018). Social rewards: From basic social building blocks to complex social behavior. *Perspectives on Psychological Science, 13*(6), 700–717.

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences, 1124*(1), 208–224.

Todorov, A. (2011). Evaluating faces on social dimensions. *Social neuroscience: Toward understanding the underpinnings of the social mind, 1*, 54–76.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66*, 519–545.

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

Vrana, S. R., & Gross, D. (2004). Reactions to facial expressions: effects of social context and speech anxiety on responses to neutral, anger, and joy expressions. *Biological Psychology, 66*(1), 63–78. https://doi.org/10.1016/j.biopsycho.2003.07.004

Zebrowitz, L. (2018). Reading faces: Window to the soul?. Routledge.