Research Article

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
1–14

DOI: 10.1177/0956797620904157
www.psychologicalscience.org/PS
SAGE

Aggressive Realism: More Efficient Processing of Anger in Physically Aggressive Individuals

Grace M. Brennan and Arielle R. Baskin-Sommers
Department of Psychology, Yale University

Abstract

Physically aggressive individuals' heightened tendency to decide that ambiguous faces are angry is thought to contribute to their destructive interpersonal behavior. Although this tendency is commonly attributed to bias, other cognitive processes could account for the emotion-identification patterns observed in physical aggression. Diffusion modeling is a valuable tool for parsing the contributions of several cognitive processes known to influence decision-making, including bias, drift rate (efficiency of information accumulation), and threshold separation (extent of information accumulation). In a sample of 90 incarcerated men, we applied diffusion modeling to an emotion-identification task. Physical aggression was positively associated with drift rate (i.e., more efficient information accumulation) for anger, and drift rate mediated the association between physical aggression and heightened anger identification. Physical aggression was not, however, associated with bias or threshold separation. These findings implicate processing efficiency for anger-related information as a potential mechanism driving aberrant emotion identification in physical aggression.

Keywords

aggressive behavior, facial expressions, face perception, cognitive processes, decision-making

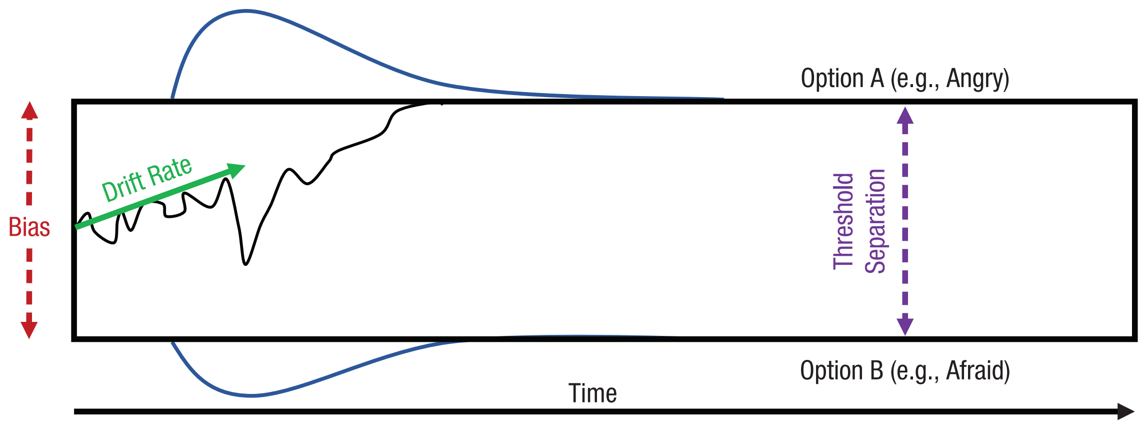
Received 7/16/19; Revision accepted 12/15/19

Physical aggression is a harmful yet ubiquitous form of human behavior. Excessive physical aggression is associated with pervasive psychosocial impairments, including low-quality friendships, social rejection, marital discord, and involvement in the criminal justice system (Bierman & Wargo, 1995; Huesmann, Dubow, & Boxer, 2009; Poulin & Boivin, 1999). Decades of research suggest that physically aggressive behavior and its associated impairments arise, in part, from a pattern of interpreting social information in aberrant ways (Crick & Dodge, 1994).

One of the richest sources of social information is facial emotion (Marsh, Ambady, & Kleck, 2005). Substantial evidence indicates that physical aggression is associated with aberrant identification of facial emotions, most notably a heightened tendency to identify ambiguous faces as angry (Mellentin, Dervisevic, Stenager, Pilegaard, & Kirk, 2015; Schönenberg & Jusyte, 2014; Wilkowski & Robinson, 2012). This *anger-perception bias*, a term that denotes impaired emotion identification and a preexisting inclination to make a particular interpretation irrespective of facial information, is theorized to drive physically aggressive behavior by fueling impressions of other individuals as hostile and threatening (Penton-Voak et al., 2013).

However, from a decision-making perspective, a response pattern observed at the behavioral level (i.e., a higher proportion of faces identified as angry) could arise from multiple cognitive processes, where "bias" is only one candidate (Ratcliff & McKoon, 2008). Because no research has looked beyond simple behavioral measures—for example, reaction time (RT) and accuracy—there has been no formal testing of the bias account, and the contributions of additional decision-making processes are unknown. Thus, although aberrant

Corresponding Author:
Grace M. Brennan, Yale University, Department of Psychology, P.O. Box 208205, New Haven, CT 06520
E-mail: grace.m.brennan@yale.edu



Parameter	Definition	Interpretation of Values
Bias	Decision Starting Point	Values Deviating From 0.5 Reflect a Bias Toward Option A (if > 0.5) or Toward Option B (if < 0.5)
Drift Rate	Average Slope of the Information-Accumulation Process	Larger Values Indicate More Efficient Processing
Threshold Separation	Amount of Information Accumulated for a Decision	Smaller Values Reflect a More Impulsive Decision-Making Style (i.e., Quicker, Less Accurate)

Fig. 1. Schematic representation of the diffusion model of decision-making. The decision process begins at a starting point that may represent an a priori response bias toward either option A or option B. Information is accumulated in favor of either option A or option B, and drift rate represents the average rate of information accumulation. The amount of information accumulated for a decision is represented by threshold separation, the distance between the two option thresholds. The "noisy" black line represents the information-accumulation process, whereas the blue curved lines above option A and below option B represent the reaction time distribution associated with each response. Diffusion-modeling parameter estimation is based on these reaction time distributions. Further information about each of the parameters shown here is given below the graph. An additional parameter estimated by diffusion modeling (not directly examined in the present study and not illustrated here) is nondecision time, which represents the length of time taken for nondecision-related processes (e.g., encoding, motor execution).

emotion identification in physical aggression is a reliable phenomenon, the underlying cognitive processes remain poorly understood.

Diffusion modeling is a form of computational modeling rooted in decision-making theory (Ratcliff, 1978) that can elucidate the cognitive processes involved in physically aggressive individuals' anger-identification patterns (Voss, Voss, & Lerche, 2015). Diffusion modeling is based on the premise that decisions are made by accumulating information until one of two response thresholds is reached, at which point the corresponding response is made (see Fig. 1). Within diffusion modeling, several processes could contribute to observed patterns of emotion identification in physical aggression. First, bias (the starting point of the decision-making process) could explain observed patterns if physically aggressive individuals require less information to identify faces as angry compared with other emotions (i.e., show a bias toward anger), predisposing them to identify faces as angry in a stimulus-nondependent manner. Second, drift rate (the rate at which information is accumulated) could explain observed patterns if physically aggressive individuals accumulate anger-related information more efficiently (i.e., show a higher drift rate for anger), leading them to identify faces as angry more swiftly without a decrement in accuracy. Finally, threshold separation (the amount of information accumulated for a decision) could contribute to observed patterns if physically aggressive individuals accumulate less information when identifying facial emotions (i.e., exhibit lower threshold separation), speeding up their responses and reducing their accuracy. Finally, any combination of these factors (e.g., lower threshold separation plus a bias toward anger) could result in an even greater likelihood of identifying faces as angry.

In addition to unspecified contributions of various cognitive processes, the impact of contextual factors (e.g., apparent motion, background scene) on emotion identification in physical aggression is unknown. Previous research indicates that contextual factors influence aggressive individuals' interpretations of social information more broadly. For example, aggressive individuals are more likely to make hostile interpretations of another's actions under conditions of threat (Dodge & Somberg, 1987). However, the impact of contextual factors on physically aggressive individuals' emotion identification has not been examined. It is possible that contextual threat amplifies physically aggressive individuals' aberrant emotion identification via altered cognitive processes (e.g., more efficient anger processing under threat, more impulsive responding under threat). Thus, in addition to quantifying cognitive processes underlying emotion identification, it is important to examine potential contributions of context to aberrant emotion identification and related cognitive processes in physical aggression.

The primary aim of the present study was to apply diffusion modeling to estimate bias, drift rate, and threshold separation during a facial emotion-identification task to test whether any of these cognitive processes could account for the association between physical aggression and aberrant emotion identification (i.e., heightened anger identification). On the basis of evidence linking physical aggression to heightened anger identification (Mellentin et al., 2015; Schönenberg & Jusyte, 2014), we hypothesized that physical aggression would be associated with a higher likelihood of identifying faces as angry (Hypothesis 1).

Additionally, despite pervasive characterizations of emotion processing as "biased" in aggression, recent findings suggest that physically aggressive individuals possess superior anger-identification abilities, exemplified by a heightened ability to discriminate between faces displaying different degrees of anger and an advanced capacity for extracting anger-related information from ambiguous faces (Wilkowski & Robinson, 2012). Rather than displaying signs of bias (i.e., showing a stimulus-nondependent tendency to identify faces as angry) or low threshold separation (i.e., responding less accurately), physically aggressive individuals display anger-identification patterns that may be most consistent with higher drift rate for anger because they appear to more efficiently and effectively accumulate information from subtle anger cues. Thus, we hypothesized that physical aggression would be associated with higher drift rate for anger (Hypothesis 2). Further, we hypothesized that drift rate would mediate the association between physical aggression and heightened anger identification (Hypothesis 3).

Finally, following research indicating that rapidly encroaching stimuli are perceived as more threatening (Coker-Appiah et al., 2013; Vieira, Tavares, Marsh, & Mitchell, 2017), we manipulated apparent movement of faces (by presenting looming or receding faces) to examine influences of contextual threat. We hypothesized that physical aggression would be associated with a higher likelihood of identifying looming (i.e., threatening) faces as angry (Hypothesis 4). Additionally, on the basis of evidence that aggression is associated with hyperreactivity to threat (Coccaro, McCloskey, Fitzgerald, & Phan, 2007; da Cunha-Bang et al., 2017) and impulsive decision-making under threat (Brennan & Baskin-Sommers, 2019; Verona & Bozzay, 2017), we hypothesized that physical aggression would be associated with lower threshold separation (i.e., greater impulsivity) under this condition (Hypothesis 5).

Method

Participants

Participants were men from a high-security correctional institution in Connecticut (for sample characteristics and correlations among key study variables, see Table S1 in the Supplemental Material available online); 96.94% of participants had been charged with a violent crime in their lifetime, and 56.12% had been charged with a violent institutional infraction while incarcerated (i.e., violations against persons, including fighting and assault on correctional staff). Because physical aggression is more pronounced in men compared with women, and because more than half of state inmates in the United States are currently serving sentences for violent crimes (Bronson & Carson, 2019), incarcerated men represent an ideal population for studying physical aggression.

Prior to recruitment, study personnel received an institutional roster of inmates. Study personnel used this roster to review medical files and exclude individuals who had a history of psychosis or bipolar disorder, currently had mood or anxiety disorders, currently used psychotropic medication, had a family history of psychosis, had medical problems that could impede comprehension of or performance on the task (e.g., uncorrectable auditory or visual deficits, three or more serious head injuries), had an IQ below 70, or had a reading level below fourth grade.

Then, individuals were selected randomly from the list of eligible inmates and invited to participate. Invited individuals were provided with information about study procedures and informed that any information collected during the study would remain confidential and would not affect their institutional or legal status in any way. They were informed that they could withdraw from the study at any time. All participants provided written informed consent. In keeping with Connecticut Department of Correction regulations, participants did not receive financial compensation. After providing consent, participants completed an initial session that involved a series of clinical and neuropsychological assessments. Participants who did not meet eligibility thresholds (detailed above) on any of these assessments were excluded from further participation. Eligible participants returned for a second session in which they completed the task followed by the aggression and emotional experience measures (see Measures section below). Both in-person sessions took place in a private testing space within the prison.

An a priori power analysis based on published studies on related topics (i.e., individual differences in facial emotion identification; Wilkowski & Robinson, 2012) indicated that a sample size of approximately 90 participants would be sufficient to detect moderate effects with 80% power. To ensure sufficient power to account for the normative loss of data due to invalid task performance, we collected data from 98 participants.

Measures

Buss-Perry Aggression Questionnaire (BPAQ). The BPAQ (Buss & Perry, 1992) is a 29-item self-report measure of aggression. Participants rate each item on a 5-point Likert-type scale (1 = *extremely uncharacteristic of me*, 5 = *extremely characteristic of me*). The four widely used subscales of the questionnaire, established through factor analysis, are Physical Aggression (9 items), Verbal Aggression (5 items), Anger (7 items), and Hostility (8 items). The BPAQ is a reliable, valid, and widely used measure of aggression (Harris, 1997; Tremblay & Ewart, 2005) and shows evidence of adequate reliability and validity in incarcerated samples (Archer & Haigh, 1997; Ireland & Archer, 2004). In the present study, we used the BPAQ Physical Aggression subscale as the measure of physical aggression, and our hypotheses centered on physical aggression on the basis of previous research (e.g., Wilkowski & Robinson, 2012). However, aggression is a multifaceted construct that can be conceptualized as having behavioral (i.e., Physical Aggression, Verbal Aggression), affective (i.e., Anger), and attitudinal (i.e., Hostility) components (Buss & Perry, 1992). Therefore, we also examined associations between these other aggression-related constructs and task performance (for analyses with the other BPAQ subscales and BPAQ total score, see the Supplemental Material). Scores for the Physical Aggression subscale can range from 5 to 45, with higher scores indicating higher levels of physical aggressiveness. Internal consistencies for the Physical Aggression subscale (Cronbach's $\alpha = .79$) and the BPAQ as a whole (Cronbach's $\alpha = .88$) in the present sample were acceptable and comparable with reliability coefficients reported by Buss and Perry (1992).

Range and Differentiation of Emotional Experience Scale (RDEES). The RDEES (Kang & Shaver, 2004) is a 14-item self-report measure of the extent to which one's emotional experiences are broad in range and well differentiated. Participants rate each item on a 5-point Likert-type scale (1 = *does not describe me very well*, 5 = *describes me very well*). The measure consists of two subscales: Range (7 items; sample item: "I experience a wide range of emotions") and Differentiation (7 items; sample item: "I am aware that each emotion has a completely different meaning"). Scores on each subscale can range from 7 to 35, with higher scores indicating greater range and differentiation of emotional experiences, respectively. Internal consistencies for both the Range and Differentiation subscales (Cronbach's αs = .67 and .80, respectively) in the present sample were acceptable. Following previous research indicating that emotion differentiation is associated with emotion-identification accuracy (Israelashvili, Oosterwijk, Sauter, & Fischer, 2019), we evaluated the validity of the task using both the Differentiation (convergent validity) and Range (divergent validity) subscales.

Ambiguous emotion-identification task. Participants completed a two-alternative forced-choice task in which they identified the emotion displayed in a series of ambiguous emotional faces.

Stimuli. Stimuli consisted of emotional face images from the Racially Diverse Affective Expression (RADIATE) face stimulus set (publicly available at http://fablab.yale.edu/page/assays-tools; Conley et al., 2018; Tottenham et al., 2009). Images of 39 unique male models of three racial/ethnic backgrounds (Black, White, and Hispanic) displaying anger, fear, and happiness were selected from the RADIATE set. The racial/ethnic composition of the face stimuli (i.e., 38.46% Black, 33.33% White, 28.21% Hispanic) roughly mirrored that found in our sample. Stimuli were generated by blending two images using face-morphing software (Abrosoft, 2018, FantaMorph Deluxe for Mac, Version 5.5.0) to create 70%–30% blends. The 70%–30% level of blending was chosen to achieve a moderate level of ambiguity (Schönenberg & Jusyte, 2014) and elicit variable but sufficiently high accuracy levels to provide data suitable for diffusion modeling (Ratcliff & McKoon, 2008). Three types of emotion blends were created: anger–fear blends, anger–happiness blends, and fear–happiness blends. We chose anger, fear, and happiness to maximize consistency with previous studies that examined emotion identification in physical aggression—outside of anger (the primary emotion of interest in the present study), fear and happiness are the most frequently used negative and positive emotions, respectively (e.g., Schönenberg & Jusyte, 2014; Wilkowski & Robinson, 2012). Within each blend, one of the two emotions served as the dominant emotion. In total, six blends per model were created (3 emotion-blend types × 2 dominant-emotion types; see Fig. 2). The process of generating six different image types for each model resulted in 234 unique images.

The task consisted of three separate blocks: an anger–fear block, an anger–happiness block, and a

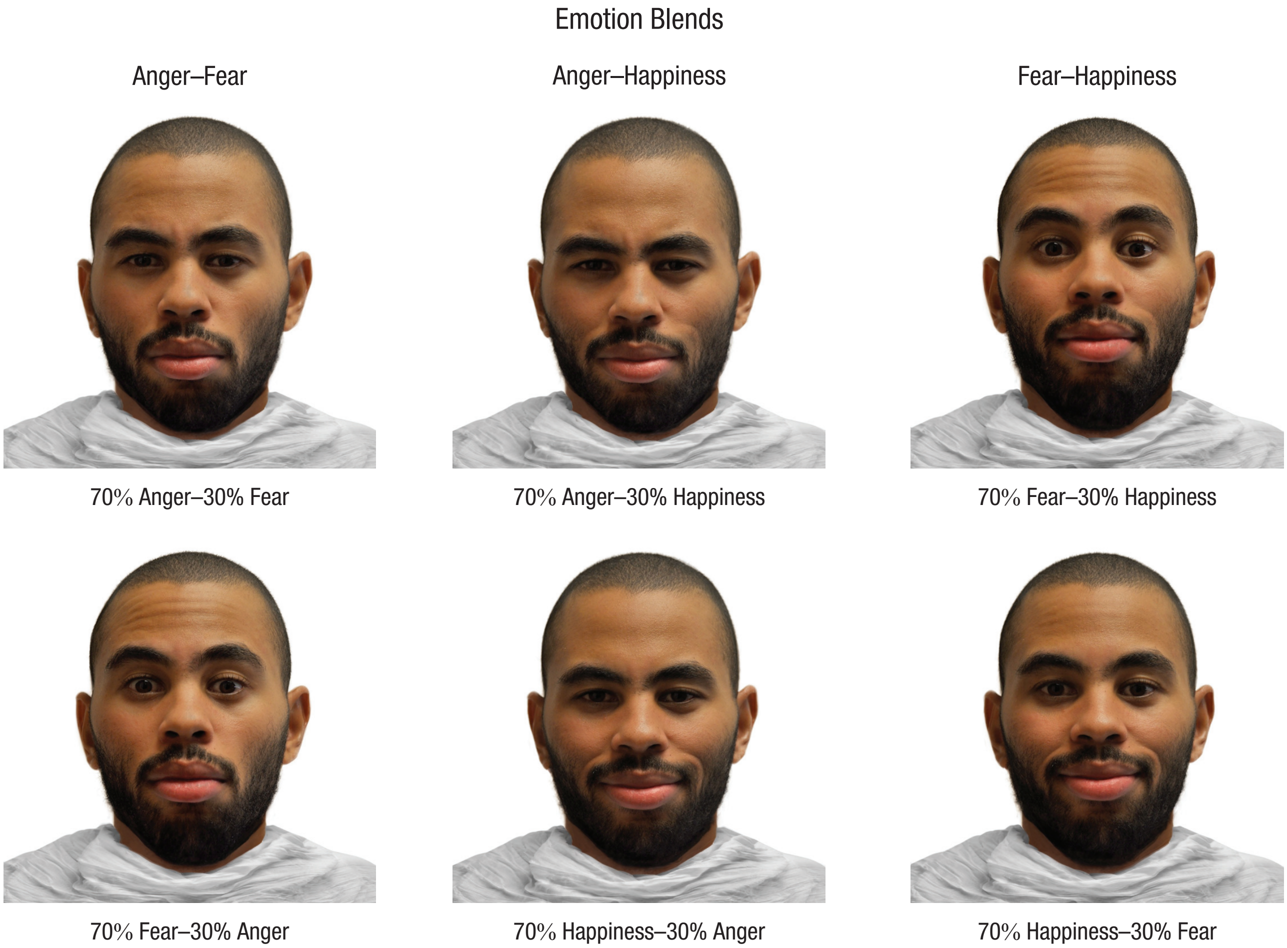


Fig. 2. Sample task stimuli. Stimuli displayed blends of anger–fear (left column), anger–happiness (middle column), and fear–happiness (right column). Within each blend type, one of the two emotions was the dominant (i.e., 70%) emotion.

fear–happiness block. Within each block, faces displayed blends of only two emotions. For example, in the anger–fear block, all faces displayed a blend of anger and fear. Because the study aims and hypotheses revolved around anger, the anger–fear and anger–happiness blocks were the primary blocks of interest; however, we included a fear–happiness block in order to examine identification patterns and decision-making parameters for fear and happiness outside of the context of anger. Within each block, half of the faces displayed mostly one emotion (e.g., 70% anger–30% fear), and half of the faces displayed mostly the other emotion (e.g., 70% fear–30% anger). Ordering of blocks was counterbalanced. Furthermore, within each block, half of the faces appeared to loom (i.e., move toward the participant), and half of the faces appeared to recede (i.e., move away from the participant). Each block consisted of 156 trials (39 unique faces × 2 dominant emotion types × 2 movement types) for a total of 468 trials in the task.

Task procedure. Participants were seated approximately 60 cm away from a 27-in. high-performance LED gaming monitor (Model XL2720Z; BenQ, Taipei, Taiwan). Participants were instructed to identify the emotion expressed in each face as quickly and accurately as possible using two keys on the keyboard. At the beginning of each block, one key was assigned to one of the two emotions represented in the faces, and another key was assigned to the other of the two emotions represented in the faces. Participants were told to press the left shift key to identify the face as one of the two emotions for that block and to press the right shift key to identify the face as the other emotion. Keyboard covers with corresponding labels were placed over the keyboard in each block to aid the participant in key–response mappings. Key–response mappings were counterbalanced across participants to counteract any effects of assigning a particular response option to either the dominant or nondominant hand. Before each block began, participants completed 10 practice trials in which they pressed the

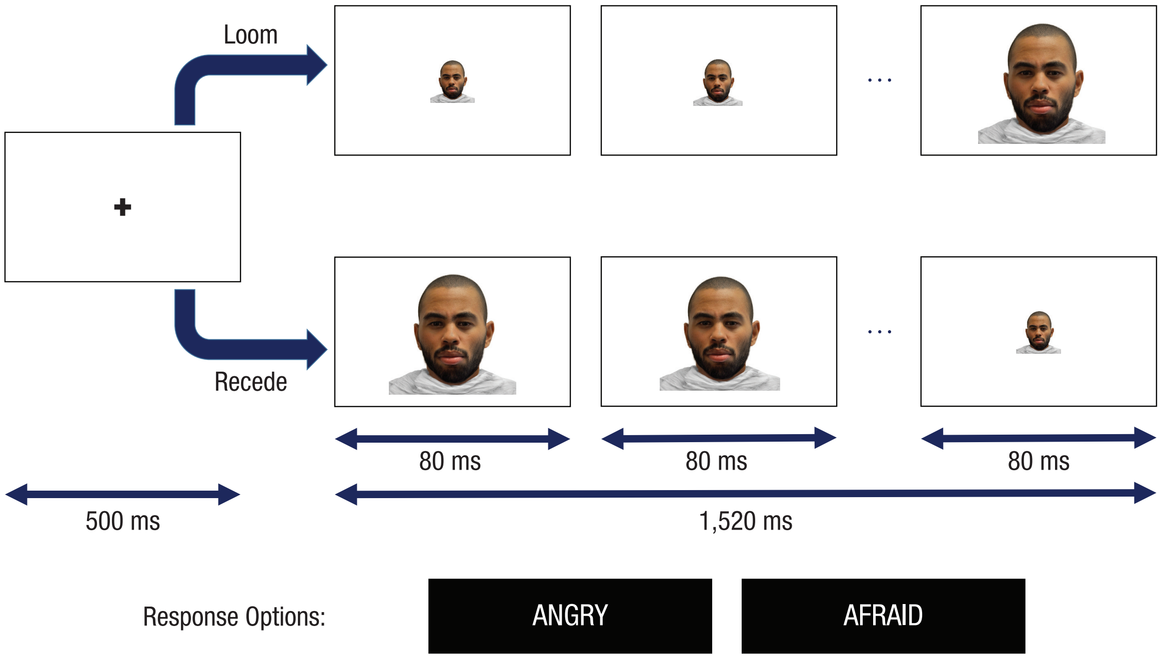


Fig. 3. Schematic representation of trial layout and timing in the ambiguous facial emotion-identification task. On each trial, participants viewed a serial presentation of images that either increased in visual angle (i.e., a looming trial, depicted in the upper row of images) or decreased in visual angle (i.e., a receding trial, depicted in the lower row of images). Shown here are the first frame, second frame, and last frame (of 19 total frames on each trial) for each trial type. Participants pressed one of two keys to identify the emotion displayed in the face.

corresponding key for the emotion word (e.g., "angry" or "afraid" prior to the anger–fear block) that appeared on the screen. To proceed to the next practice trial (and ultimately to the main task), participants were required to press the correct key on each practice trial (and were given multiple chances, if needed).

Stimulus presentation and response collection were controlled using the Psychophysics Toolbox extension (Version 3; Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997) as implemented in MATLAB (The MathWorks, Natick, MA). Stimuli were presented in random order for each participant. Each trial began with a fixation cross (500 ms), after which a face was displayed on the screen for a total of 1,520 ms. Following previous research (Vieira et al., 2017), we created movement effects by rapidly changing the visual angle of stimuli. Faces increased (on looming trials) or decreased (on receding trials) in size by a factor of 1.05, resulting in 19 frames (each lasting 80 ms) per trial (see Fig. 3). The intertrial interval varied randomly between 1,000 and 2,000 ms (average = 1,500 ms).

Control emotion-identification task. After completing the main task, participants completed a control emotion-identification task, which assessed general emotion-identification accuracy as a way to examine the validity of the ambiguous emotion-identification task. Participants were instructed to choose the emotion displayed by each face that appeared on the screen. Stimuli in the control task consisted of unblended emotional face images (i.e., 100% display of one emotion) from the RADIATE stimulus set. The emotions displayed in the images were the same as those used in the main task: anger, fear, and happiness. Participants were given three response options (one for each emotion), which appeared as text in three separate panels below each face on the screen, and they used a mouse to click the panel corresponding to the emotion that each face displayed. There was no time limit imposed for responding. The control task consisted of 54 trials.

Data processing and analysis

Data quality control. Participants were excluded from analyses if their task data were invalid. Data were considered invalid if at least one of the following conditions was met: (a) no response given (or response given in < 300 ms) on more than 20% of trials, (b) accuracy at or below chance (i.e., $\leq 50\%$), or (c) statistical outliers (> 3 *SD*s from the mean) on task behavioral variables. Seven participants were excluded from the analyses on the basis of these criteria, and the resulting sample consisted of 91 participants.

Task validation. Convergent validity of the ambiguous emotion-identification task was evaluated by examining associations between the RDEES Differentiation score and overall task accuracy, as well as between control task accuracy and overall task accuracy. Divergent validity of the task was evaluated by examining the association between RDEES Range score and overall task accuracy. On the one hand, we expected RDEES Differentiation (one's ability to identify subtle variations in emotional experiences) to be positively related to task accuracy (i.e., an enhanced ability to correctly identify the dominant emotion in ambiguous emotional faces; Israelashvili et al., 2019), and we also expected control task accuracy to be positively related to task accuracy, because the control task had similar demands but no stimulus movement effects and no time limit for responses. On the other hand, we did not expect RDEES Range (one's own experience of a range of different emotions) to be related to task accuracy.

A 3 (emotion blend: anger–fear, anger–happiness, fear–happiness) × 2 (movement: looming, receding) repeated measures general linear model (GLM), with RDEES Differentiation (z scored) as a continuous between-subjects independent variable and overall task accuracy as a dependent variable, revealed a main effect of differentiation on accuracy, $F(1, 88) = 4.19$, $p = .044$, $\eta_p^2 = .05$, 90% confidence interval (CI) = [.001, .13]; individuals with higher levels of differentiation exhibited higher emotion-identification accuracy overall ($b = 3.61$, $SE = 1.76$, 95% CI = [0.10, 7.11]), providing support for convergent validity of the task. Furthermore, we detected a moderately strong correlation between control task accuracy and overall task accuracy, $r(88) = .49$, providing additional support for convergent validity of the task (see Table S1). A 3 (emotion blend: anger–fear, anger–happiness, fear–happiness) × 2 (movement: looming, receding) repeated measures GLM, with RDEES Range (z scored) as a continuous between-subjects independent variable and overall task accuracy as a dependent variable, revealed no associations between RDEES Range and overall task accuracy, providing support for divergent validity of the task.

Diffusion modeling. Following established guidelines (Voss et al., 2015), we removed trials with no response (i.e., omissions) and trials with RTs less than 300 ms (i.e., premature responses) from individual participants' data before subjecting them to diffusion modeling. Rates of omissions and premature responses were low (i.e., omissions characterized, on average, 2.51% of trials per participant, and premature responses characterized 0.24% of trials).

We used fast-dm-30 software (Voss & Voss, 2007; Voss et al., 2015) to estimate decision-making parameters on the basis of response and RT data from the task. The software was designed to estimate parameters from Ratcliff's (1978) diffusion model, in which decision-making is a noisy, continuous process of accumulating information until one of two decision thresholds (one for each response option) is met. The model uses RT distributions for the two response options to estimate decision-making parameters, including bias, threshold separation, and drift rate. By using the entire range of task performance across trials (rather than simple accuracy or mean RTs in isolation), diffusion modeling delivers several advantages over traditional methods, including increased reliability and the potential to yield novel mechanistic insights (Price, Brown, & Siegle, 2019). The Kolmogorov-Smirnov estimation procedure was used because it accounts for exact RT distributions (as opposed to binning RT data) and is robust to contaminants. Guided by theoretical and methodological considerations, we allowed the relative starting point to vary by emotion blend (i.e., block) only (because starting point is not impacted by stimulus features), and we allowed both threshold separation and drift rate to vary by all three conditions (i.e., emotion blend, dominant emotion, and movement). "Angry" responses were set as response option A, whereas non-"angry" responses were set as response option B (see Fig. 1), so that positive bias-parameter values would indicate a bias toward anger, and positive drift-rate values would indicate a drift rate toward anger (and conversely, negative values would indicate a bias toward the nonanger emotion—i.e., happiness or fear, depending on the block—and drift rate toward the nonanger emotion, respectively). Parameter values for threshold separation, by contrast, are not directional, and they typically range from 0.5 to 2.0. To maximize parsimony and accuracy of the model, we opted for a four-parameter model, in which our three parameters of interest plus nondecision time were allowed to vary, whereas the remaining parameters were fixed at 0 (Lerche & Voss, 2016; for correlations among the diffusion-modeling parameters, see Table S2 in the Supplemental Material).

Here, we provide a simplified illustration of how various patterns of RT distributions impact the diffusion-modeling parameter estimates. In a two-choice task, each response option (e.g., option A and option B in Fig. 1) has an RT distribution that is determined by the frequency of different RTs for that response across all trials within a given task condition. The frequency of RTs for one response determines the RT distribution for that response option (e.g., the blue curved line above option A in Fig. 1), whereas the frequency of RTs for the other response determines the RT distribution for that response option (e.g., the blue curved line below option B in Fig. 1). For example, if a participant has a large number of fast RTs and virtually no slow RTs when making one response, the RT distribution for that response option will be compressed toward the left (i.e., fast) end of the distribution. In terms of how various RT distribution patterns translate into diffusion-modeling parameter estimates, let us consider which RT distribution patterns would correspond to a bias toward one response option, higher drift rate toward one response option, and lower threshold separation. For bias, if the leading end of the RT distribution for option A is compressed toward the left without a corresponding compression in the RT distribution for option B, then diffusion modeling estimates a higher value for the bias parameter (i.e., greater than 0.5), indicating a bias toward option A. For drift rate, if there is an increased relative probability of faster RTs for option A (i.e., the RT distribution is taller for option A), then diffusion modeling estimates a higher value for the drift-rate parameter, indicating stronger drift rate toward option A. Finally, for threshold separation, if the RT distributions for both response options are compressed toward the left, then diffusion modeling estimates a lower value for the threshold-separation parameter. Although both bias and lowered threshold separation are associated with reduced performance on a task, each parameter relates to performance in a distinct manner. Whereas a bias toward one response option would preferentially increase the likelihood of that response, lowered threshold separation, which denotes the extent of information accumulation for both response options, would decrease accuracy in general but not in favor of either response option.

Following parameter estimation, model fit was assessed using Kolmogorov-Smirnov test statistics (values > .05 generally indicate acceptable fit), along with visual inspection of quantile–quantile (Q-Q) plots (which indicate acceptable fit if all data points lie near the main diagonal). These indices revealed that the model generally fitted the data well. On the basis of visual inspection, we deemed that one participant's data fitted poorly to the model, and this participant was excluded. Thus, the final sample consisted of 90 participants. Excluded participants did not differ from included participants in terms of physical aggression (95% CI for the mean difference = [−5.40, 5.78], p = .905).

The study protocol was approved by the Yale University Human Investigation Committee and was carried out in accordance with the provisions of the World Medical Association Declaration of Helsinki. In this article, we report all of the dependent measures collected, all data exclusions, and all of the task conditions. This study was not preregistered. The data have not been made available on a permanent third-party archive because the combination of demographic and crime variables makes it possible to identify participants. However, requests for a deidentified subset of the data can be e-mailed to the corresponding author.

Results

Emotion identification

Given previous research highlighting heightened anger identification in physical aggression, we started by conducting a 2 (emotion blend: anger–fear, anger–happiness) × 2 (dominant emotion: anger, nonanger) × 2 (movement: looming, receding) repeated measures GLM with BPAQ Physical Aggression (*z* scored) as a continuous between-subjects independent variable and the proportion of trials on which participants responded "angry" (i.e., anger identification) as a dependent variable (for additional analyses pertaining to robustness of results and specificity to physical aggression, see the Supplemental Material). The analysis revealed both task effects and physical-aggression-related effects.

In terms of task effects, there was a main effect of dominant emotion on anger identification, $F(1, 88) = 1845.92$, $p < .001$, $\eta_p^2 = .95$, 90% CI = [.94, .96]; mostly angry faces were more likely to be identified as angry (M = 73.7%, 95% CI = [71.8%, 75.6%]) compared with mostly nonangry faces (M = 21.6%, 95% CI = [20.0%, 23.2%]). This main effect provides a key demonstration of task validity by indicating that participants were able to discriminate between the two types of faces and identify the dominant emotion.

There was also a main effect of movement on anger identification, $F(1, 88) = 5.16$, $p = .025$, $\eta_p^2 = .10$, 90% CI = [.004, .15]; looming faces were more likely to be identified as angry (M = 48.2%, 95% CI = [46.8%, 49.6%]) compared with receding faces (M = 47.1%, 95% CI = [45.8%, 48.5%]). This effect provides support for the success of the movement manipulation and suggests that looming faces were perceived as more threatening. Additionally, there was an Emotion Blend × Movement interaction, $F(1, 88) = 9.78$, $p = .002$, $\eta_p^2 = .06$, 90% CI = [.02, .20]; looming faces were more likely to be identified as angry, and this was particularly true for the anger–fear blended faces (M = 49.1% for looming faces, 95% CI = [47.2%, 51.0%]; M = 46.4% for receding faces, 95% CI = [44.3%, 48.6%]) compared with the anger–happiness blended faces (M = 47.3% for looming faces, 95% CI = [45.6%, 49.0%]; M = 47.8% for receding faces, 95% CI = [46.3%, 49.4%]). An Emotion Blend × Dominant Emotion interaction emerged as well, $F(1, 88) = 293.81$, $p < .001$, $\eta_p^2 = .77$, 90% CI = [.70, .81], indicating that the difference in anger identification as a function of the dominant emotion in the face was greater for the anger–happiness blended faces (M = 79.5% for mostly angry faces, 95% CI = [77.7%, 81.3%];

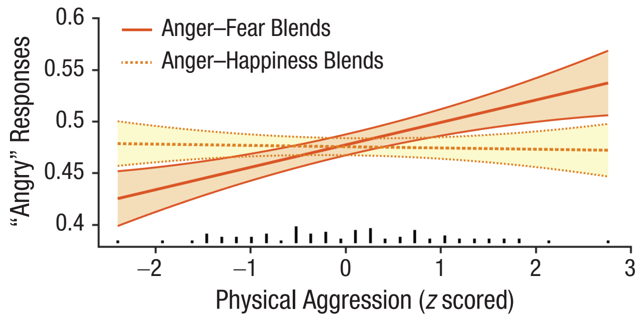


Fig. 4. The relationship between physical aggression and proportion of "angry" responses for anger–fear blended faces and anger–happiness blended faces. Error bands represent ±1 *SE*, and the dot plot along the *x*-axis represents frequencies for Buss-Perry Aggression Questionnaire Physical Aggression scores.

M = 15.6% for mostly nonangry faces, 95% CI = [13.7%, 17.6%]) compared with the anger–fear blended faces (M = 67.9% for mostly angry faces, 95% CI = [65.0%, 70.8%]; M = 27.6% for mostly nonangry faces, 95% CI = [25.6%, 29.6%]), suggesting that participants had greater difficulty accurately identifying the dominant emotion for faces that displayed a blend of anger and fear. This finding indicates that the anger–fear blended faces were significantly more ambiguous than the anger–happiness blended faces.

In terms of physical-aggression-related effects, there was an Emotion Blend × Physical Aggression interaction, $F(1, 88) = 4.07$, $p = .047$, $\eta_p^2 = .04$, 90% CI = [.0003, .13]; physical aggression was positively related to the proportion of "angry" responses for the anger–fear blended faces ($b = 0.02$, $SE = 0.01$, 90% CI = [0.01, 0.04], $p = .025$, $\eta_p^2 = .06$), but no association was detected for the anger–happiness blended faces ($b = -0.01$, $SE = 0.01$, 90% CI = [−0.01, 0.01], $p = .873$, $\eta_p^2 = .00$; see Fig. 4). Results remained unchanged after we controlled for participant race. Thus, the results were consistent with Hypothesis 1: Higher levels of physical aggression were associated with a heightened tendency to identify ambiguous faces as angry. However, we failed to find support for Hypothesis 4, as there was no interaction involving movement and physical aggression.

Because physical aggression was associated with heightened anger identification for the anger–fear blended faces (but not the anger–happiness blended faces), it is possible that this association was confounded by fear identification. That is, if individuals with higher levels of physical aggression were generally less likely to identify faces as afraid, this could have accounted for their heightened tendency to identify faces as angry when they were presented with two response options: angry or afraid. To rule out fear identification as a potential confound of the association between physical aggression and anger identification, we analyzed fear identification outside of the context of anger (i.e., in the fear–happiness block) by conducting a 2 (dominant emotion: fear, happiness) × 2 (movement: looming, receding) repeated measures GLM with BPAQ Physical Aggression (z scored) as a continuous between-subjects independent variable and proportion of "afraid" responses as a dependent variable. The GLM revealed a significant main effect of dominant emotion on fear identification, $F(1, 88) = 2940.83$, $p < .001$, $\eta_p^2 = .97$, 90% CI = [.96, .98]; mostly afraid faces were more likely to be identified as afraid (M = 83.3%, 95% CI = [81.6%, 85.0%]) compared with mostly happy faces (M = 19.1%, 95% CI = [16.9%, 21.2%]). There were no other task effects and, crucially, no physical-aggression-related effects associated with fear identification. Most notably, we failed to detect a main effect of physical aggression on fear identification, $F(1, 88) = 0.07$, $p = .786$, $\eta_p^2 = .001$, 90% CI = [.00, .03]. The failure to detect physical-aggression-related effects associated with fear identification suggests that the association between physical aggression and heightened anger identification for anger–fear blended faces is not attributable to a diminished tendency to identify fear in faces.

Diffusion-modeling parameters

To examine decision-making parameters estimated with diffusion modeling as a function of task conditions as well as physical aggression, we conducted a series of 2 (emotion blend: anger–fear, anger–happiness) × 2 (dominant emotion: anger, nonanger) × 2 (movement: looming, receding) repeated measures GLMs with BPAQ Physical Aggression (z scored) as a continuous between-subjects independent variable and each diffusion-modeling parameter as a dependent variable.

Bias. There were no task effects and no physical-aggression-related effects associated with bias (all ps ≥ .352).

Drift rate. Examination of drift rate as a dependent variable revealed both task effects and physical-aggression-related effects. In terms of task effects, there was a main effect of dominant emotion on drift rate, $F(1, 88) = 1161.87$, $p < .001$, $\eta_p^2 = .93$, 90% CI = [.91, .94], indicating that drift rate toward anger was higher for mostly angry faces (M = 0.89, 95% CI = [0.82, 0.96]) compared with mostly nonangry faces (M = −0.94, 95% CI = [−1.00, −0.87]). There was also an Emotion Blend × Movement interaction, $F(1, 88) = 8.67$, $p = .004$, $\eta_p^2 = .09$, 90% CI = [.02, .19], indicating that drift rate toward anger was higher for looming faces, but only for the anger–fear blended faces (M = 0.03 for looming faces, 95% CI = [−0.05, 0.10]; M = −0.07 for receding faces, 95% CI = [−0.15, 0.01]) and not for the anger–happiness blended faces ($M = -0.05$ for looming faces, 95% CI = [−0.13, 0.02]; $M = 0.00$ for receding faces, 95% CI = [−0.07, 0.07]). Additionally, there was an Emotion Blend × Dominant Emotion interaction, $F(1, 88) = 262.18$, $p < .001$, $\eta_p^2 = .75$, 90% CI = [.67, .80], indicating that the difference between drift rates for mostly angry and mostly nonangry faces was greater for the anger–happiness blended faces ($M = 1.14$ for mostly angry faces, 95% CI = [1.07, 1.21]; $M = -1.19$ for mostly nonangry faces, 95% CI = [−1.28, −1.10]) compared with the anger–fear blended faces ($M = 0.64$ for mostly angry faces, 95% CI = [0.54, 0.74]; $M = -0.68$ for mostly nonangry faces, 95% CI = [−0.76, −0.60]). The fact that drift rate was more strongly differentiated according to dominant emotion for the anger–happiness faces (i.e., information accumulation proceeded more efficiently) is again consistent with the idea noted above that the anger–fear blended faces were significantly more ambiguous than the anger–happiness blended faces.

The final task effect was a Movement × Dominant Emotion interaction, $F(1, 88) = 9.05$, $p = .003$, $\eta_p^2 = .09$, 90% CI = [.02, .20], indicating that the difference between drift rate for mostly angry and mostly nonangry faces was greater for receding faces ($M = 0.91$ for mostly angry faces, 95% CI = [0.84, 0.99]; $M = -0.98$ for mostly nonangry faces, 95% CI = [−1.05, −0.90]) compared with looming faces ($M = 0.87$ for mostly angry faces, 95% CI = [0.80, 0.94]; $M = -0.89$ for mostly nonangry faces, 95% CI = [−0.96, −0.82]).

In terms of physical-aggression-related effects, there was an Emotion Blend × Physical Aggression interaction, $F(1, 88) = 5.32$, $p = .023$, $\eta_p^2 = .06$, 90% CI = [.004, .15]; physical aggression was positively related to drift rate for the anger–fear blended faces ($b = 0.11$, $SE = 0.04$, 90% CI = [0.04, 0.18], $p = .014$, $\eta_p^2 = .07$), but no association was detected for the anger–happiness blended faces ($b = -0.02$, $SE = 0.04$, 90% CI = [−0.09, 0.04], $p = .599$, $\eta_p^2 = .00$; see Fig. 5). Results remained unchanged after we controlled for participant race. Thus, consistent with Hypothesis 2, higher levels of physical aggression were associated with higher drift rate for anger.

Because physical aggression was associated with higher drift rate for anger when identifying the anger–fear blended faces (but not the anger–happiness blended faces), it is possible that this association was confounded by drift rate for fear. That is, if individuals with higher levels of physical aggression had a lower drift rate for fear in general, this could have accounted for their higher drift rate for anger when they were presented with two response options: angry or afraid. To rule out drift rate for fear as a potential confound of the association between physical aggression and drift rate for anger, we analyzed drift rate for fear outside of the context of anger (i.e., in the fear–happiness block)

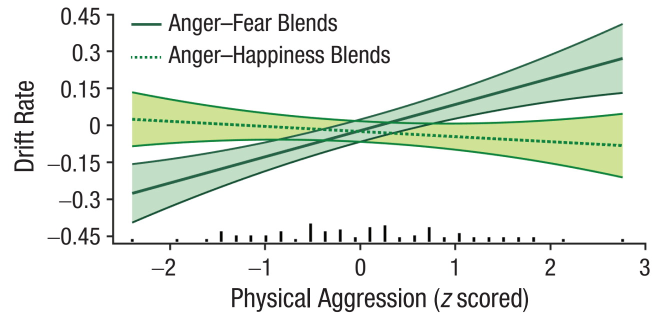


Fig. 5. The relationship between physical aggression and drift rate toward anger for anger–fear blended faces and anger–happiness blended faces. Error bands represent ±1 *SE*, and the dot plot along the *x*-axis represents frequencies for Buss-Perry Aggression Questionnaire Physical Aggression scores.

by conducting a 2 (dominant emotion: fear, happiness) × 2 (movement: looming, receding) repeated measures GLM with BPAQ Physical Aggression (z scored) as a continuous between-subjects independent variable and drift rate for fear as a dependent variable. The GLM revealed a significant main effect of dominant emotion on drift rate for fear, $F(1, 88) = 1426.22$, $p < .001$, $\eta_p^2 = .94$, 90% CI = [.92, .95]; drift rate for fear was higher for mostly afraid faces ($M = 1.38$, 95% CI = [1.29, 1.47]) compared with mostly happy faces ($M = -1.07$, 95% CI = [−1.18, −0.96]). There were no other task effects and, crucially, no physical-aggression-related effects associated with drift rate for fear. Most notably, we failed to detect a main effect of physical aggression on drift rate for fear, $F(1, 88) = 0.14$, $p = .714$, $\eta_p^2 = .00$, 90% CI = [.00, .04]. The failure to detect physical-aggression-related effects associated with drift rate for fear suggests that the association between physical aggression and higher drift rate for anger when identifying anger–fear blended faces is not attributable to less efficient processing of fear-related information.

Threshold separation. Examination of threshold separation as a dependent variable revealed task effects but no physical-aggression-related effects (indicating a failure to find support for Hypothesis 5). In terms of task effects, there was a main effect of emotion blend, $F(1, 88) = 34.57$, $p < .001$, $\eta_p^2 = .28$, 90% CI = [.16, .39], indicating that threshold separation was greater for the anger–happiness blended faces ($M = 1.25$, 95% CI = [1.23, 1.27]) compared with the anger–fear blended faces ($M = 1.19$, 95% CI = [1.18, 1.21]). There was also a main effect of movement, $F(1, 88) = 16.82$, $p < .001$, $\eta_p^2 = .16$, 90% CI = [.06, .27], indicating that threshold separation was lower for looming faces ($M = 1.20$, 95% CI = [1.19, 1.22]) compared with receding faces ($M = 1.24$, 95% CI = [1.22, 1.26]). Finally, there was an Emotion Blend × Dominant Emotion interaction, $F(1, 88) = 6.11$, $p = .015$, $\eta_p^2 = .07$, 90% CI = [.01, .16], indicating that threshold separation was lower for mostly angry faces but only for the anger–happiness blended faces (M = 1.24 for mostly angry faces, 95% CI = [1.21, 1.26]; M = 1.26 for mostly nonangry faces, 95% CI = [1.24, 1.29]) and not for the anger–fear blended faces (M = 1.21 for mostly angry faces, 95% CI = [1.19, 1.23]; M = 1.18 for mostly nonangry faces, 95% CI = [1.16, 1.20]).

Mediation model

Given that the goal of the present study was to identify potential mechanisms supporting the link between physical aggression and heightened anger identification, we conducted a mediation analysis with BPAQ Physical Aggression as the independent variable, proportion of "angry" responses for the anger–fear blended faces as the dependent variable, and drift rate toward anger for the anger–fear blended faces as the mediator (see Fig. 6). The analysis was performed using the PROCESS macro for SPSS (Hayes, 2018), Model 4. We used a nonparametric resampling procedure (bootstrapping) with 5,000 samples to estimate the indirect effect. The analysis indicated a significant indirect effect of physical aggression on anger identification through drift rate (b = 0.004, SE = 0.002, 95% CI = [0.001, 0.007]). Thus, drift rate mediated the association between physical aggression and heightened anger identification, consistent with Hypothesis 3.

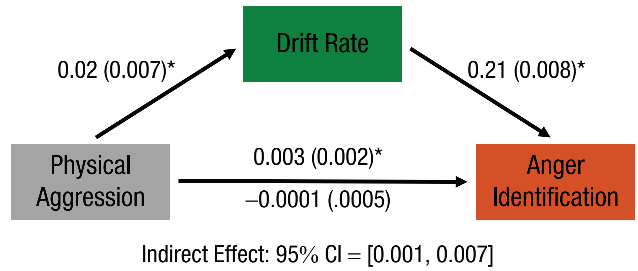


Fig. 6. Mediation model showing the influence of physical aggression on anger identification (i.e., proportion of "angry" responses) for the anger–fear blended faces, as mediated by drift rate toward anger. On the path from physical aggression to anger identification, the value above the arrow shows the total effect, and the value below the arrow shows the direct effect after controlling for the mediator. Unstandardized coefficients are shown, with standard errors in parentheses. The confidence interval (CI) for the indirect effect is also shown. Asterisks indicate significant paths (p < .05).

Discussion

Substantial evidence indicates that physically aggressive individuals exhibit a heightened tendency to identify anger in ambiguous faces. The present study was the first empirical endeavor to apply diffusion modeling to identify contributions of cognitive processes (i.e., bias, efficiency of information accumulation, extent of information accumulation) to this tendency. Results suggest that physically aggressive individuals' aberrant emotion identification (i.e., heightened anger identification for anger–fear faces) stems from more efficient processing of anger-related cues (i.e., higher drift rate) rather than from bias or impulsive responding (i.e., threshold separation). Moreover, higher drift rate mediated the association between physical aggression and heightened anger identification, highlighting the role of processing efficiency for anger-related information in physically aggressive individuals' propensity to arrive at aggression-promoting interpretations of social information.

The finding that physical aggression was associated with heightened anger identification for highly ambiguous (i.e., anger–fear) faces is consistent with previous research indicating aberrant social interpretations in aggression only under high ambiguity (Dodge, 1980; Mellentin et al., 2015; Schönenberg & Jusyte, 2014; Wilkowski & Robinson, 2012; Zimmer-Gembeck & Nesdale, 2013). Yet physical aggression was not associated with overall task accuracy (see Table S1), indicating that more physically aggressive individuals did not erroneously identify faces as angry. Indeed, physical aggression was positively correlated with accuracy for mostly angry anger–fear faces, suggesting that these individuals were *more* accurate under high ambiguity.

The present study's key contribution is demonstrating that more physically aggressive individuals appear to be more efficient at accumulating information related to anger under highly ambiguous conditions (i.e., for anger–fear faces), and this heightened efficiency may explain their tendency to see anger where less aggressive individuals do not. Although the concept of bias is inherent in the terms used to describe aggressive individuals' patterns of interpreting social information (e.g., anger-perception bias, hostile-attribution bias), our results do not support the contention that more physically aggressive individuals display impairments or biases in emotion identification. Instead, results suggest that these individuals are more adept at processing anger-related information, building on evidence that physical aggression relates to superior anger-identification abilities (Wilkowski & Robinson, 2012). Because drift rate indexes information accumulation not only from perception but also from memory (Ratcliff, Smith, Brown, & McKoon, 2016), this finding can be interpreted in light of theory positing that aggressive individuals have stronger and more accessible hostile knowledge structures, which are essentially latent memories of hostility-related events (Anderson & Bushman, 2002). That is, as aggressive individuals accrue experiences of hostile interactions (brought about in part through their own aggressive behavior; Anderson, Buckley, & Carnagey, 2008), they activate and build on their existing hostile knowledge structures, making these structures more readily accessible to aid in interpreting incoming social information. Our findings suggest, however, that rather than drawing on knowledge structures to make biased interpretations of social information, physically aggressive individuals draw on knowledge structures to make more accurate interpretations, allowing them to adopt a realistic lens for viewing their often hostile worlds.

Although we did not find physical-aggression-related effects of movement, we found that participants were generally more likely to identify looming faces as angry, and this tendency was stronger for the more ambiguous anger–fear faces. Whereas previous research has shown that looming objects and faces elicit greater threat-related neural activity (Coker-Appiah et al., 2013; Vieira et al., 2017), our findings provide the first demonstration that looming ambiguous faces are more likely to be identified as angry. Thus, threat-based reactivity to looming faces may impact actual emotion identification; individuals (regardless of level of physical aggressiveness) are more likely to "see" anger in rapidly encroaching faces. Moreover, following from the diffusion-modeling results, because threshold separation was lower for looming faces, it is possible that more impulsive responding in the context of threat leads individuals to identify anger in ambiguous faces. The failure to detect an association between physical aggression and heightened anger identification for looming faces is again inconsistent with the view that physically aggressive individuals exhibit impairments in interpreting social information, because they were no more or less likely to be misled by apparent movement, a contextual factor that was orthogonal to the emotion displayed in the faces.

Several limitations of the present study should be noted. First, because our sample was limited to male offenders, it is unclear whether the results would generalize to other populations. However, because male offenders perpetrate physical violence at high rates, understanding aggression in this population is particularly important. Future research should seek to replicate findings in female and nonincarcerated (e.g., community) samples. Second, we did not present faces of varying emotional intensities, a more direct manipulation of ambiguity. Results indicated that physical-aggression-related effects were specific to anger–fear faces rather than extending to anger–happiness as well, which may reflect the tendency among physically aggressive individuals to process anger differently only under high-ambiguity conditions. However, because we did not directly manipulate ambiguity, we cannot rule out the possibility that the physical-aggression-related effects for the more ambiguous stimuli (i.e., anger–fear faces) are attributable to the specific anger–fear blend rather than greater ambiguity per se. However, failure to find physical-aggression-related differences in general fear identification and drift rate for fear provides evidence against the interpretation that the results are an artifact of the anger–fear blend. Future research should directly manipulate ambiguity and use other types of emotion blends (e.g., anger–sadness) to test the generalizability of the present findings to other ambiguous stimuli.

Overall, the present study contributes to mounting evidence that physical aggression is associated with aberrant processing of anger. Although researchers have used the term *bias* to describe physically aggressive individuals' anger-processing aberrations, the present study suggests that their aberrant processing stems from efficiency and adeptness. Thus, physical aggression may be characterized by *aggressive realism*, or a tendency to more readily process anger when it is present in ambiguous social stimuli. Progress in understanding the mechanisms contributing to physical aggression may be made by investigating how seemingly adaptive capabilities can lead to maladaptive social behaviors.

Transparency

Action Editor: Alice J. O'Toole
Editor: D. Stephen Lindsay
Author Contributions

G. M. Brennan developed the study concept. Both G. M. Brennan and A. R. Baskin-Sommers designed the study and performed the clinical and neuropsychological assessments. G. M. Brennan collected some of the task data. G. M. Brennan analyzed and interpreted the data under the supervision of A. R. Baskin-Sommers. G. M. Brennan drafted the manuscript, and A. R. Baskin-Sommers provided critical revisions. Both authors approved the final manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported by the Harry Frank Guggenheim Foundation and by the American Psychology-Law Society, American Psychological Association Division 41.

Open Practices

Data for this study have not been made publicly available because the combination of demographic and crime variables makes it possible to identify participants. However, requests for a deidentified subset of the data can be e-mailed to the corresponding author. The face stimuli used in the present study are morphed versions of images that can be downloaded at http://fablab.yale.edu/page/assays-tools. The design and analysis plans for this study were not preregistered.

ORCID iD

Grace M. Brennan https://orcid.org/0000-0003-0246-0820

Acknowledgments

We thank the Connecticut Department of Correction staff for their support and the research assistants for helping to collect the data.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797620904157

References

Abrosoft. (2018). FantaMorph Deluxe for Mac (Version 5.5.0) [Computer software]. Retrieved from https://www.fantamorph.com/download.html

Anderson, C. A., Buckley, K. E., & Carnagey, N. L. (2008). Creating your own hostile environment: A laboratory examination of trait aggressiveness and the violence escalation cycle. *Personality and Social Psychology Bulletin*, *34*, 462–473.

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology*, *53*, 27–51. doi:10.1146/annurev.psych.53.100901.135231

Archer, J., & Haigh, A. (1997). Beliefs about aggression among male and female prisoners. *Aggressive Behavior*, *23*, 405–415. doi:10.1002/(SICI)1098-2337(1997)23:6<405::AID-AB1>3.0.CO;2-F

Bierman, K. L., & Wargo, J. B. (1995). Predicting the longitudinal course associated with aggressive-rejected, aggressive (nonrejected), and rejected (nonaggressive) status. *Development and Psychopathology*, *7*, 669–682. doi:10.1017/S0954579400006775

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Brennan, G. M., & Baskin-Sommers, A. R. (2019). Physical aggression is associated with heightened social reflection impulsivity. *Journal of Abnormal Psychology*, *128*, 404–414. doi:10.1037/abn0000448

Bronson, J., & Carson, E. A. (2019). *Prisoners in 2017* (NCJ Publication No. 252156). Retrieved from https://www.bjs.gov/content/pub/pdf/p17.pdf

Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, *63*, 452–459.

Coccaro, E. F., McCloskey, M. S., Fitzgerald, D. A., & Phan, K. L. (2007). Amygdala and orbitofrontal reactivity to social threat in individuals with impulsive aggression. *Biological Psychiatry*, *62*, 168–178. doi:10.1016/j.biopsych.2006.08.024

Coker-Appiah, D. S., White, S. F., Clanton, R., Yang, J., Martin, A., & Blair, R. J. (2013). Looming animate and inanimate threats: The response of the amygdala and periaqueductal gray. *Social Neuroscience*, *8*, 621–630. doi:10.1080/17470919.2013.839480

Conley, M. I., Dellarco, D. V., Rubien-Thomas, E., Cohen, A. O., Cervera, A., Tottenham, N., & Casey, B. J. (2018). The racially diverse affective expression (RADIATE) face stimulus set. *Psychiatry Research*, *270*, 1059–1067. doi:10.1016/j.psychres.2018.04.066

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*, 74–101.

da Cunha-Bang, S., Fisher, P. M., Hjordt, L. V., Perfalk, E., Persson Skibsted, A., Bock, C., . . . Knudsen, G. M. (2017). Violent offenders respond to provocations with high amygdala and striatal reactivity. *Social Cognitive & Affective Neuroscience*, *12*, 802–810. doi:10.1093/scan/nsx006

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development*, *51*, 162–170.

Dodge, K. A., & Somberg, D. R. (1987). Hostile attributional biases among aggressive boys are exacerbated under conditions of threats to the self. *Child Development*, *58*, 213–224. doi:10.1111/1467-8624.ep7264206

Harris, J. A. (1997). A further evaluation of the Aggression Questionnaire: Issues of validity and reliability. *Behaviour Research and Therapy*, *35*, 1047–1053. doi:10.1016/S0005-7967(97)00064-8

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). New York, NY: Guilford Press.

Huesmann, L. R., Dubow, E. F., & Boxer, P. (2009). Continuity of aggression from childhood to early adulthood as a predictor of life outcomes: Implications for the adolescent-limited and life-course-persistent models. *Aggressive Behavior*, *35*, 136–149. doi:10.1002/ab.20300

Ireland, J. L., & Archer, J. (2004). Association between measures of aggression and bullying among juvenile and young offenders. *Aggressive Behavior*, *30*, 29–42. doi:10.1002/ab.20007

Israelashvili, J., Oosterwijk, S., Sauter, D., & Fischer, A. (2019). Knowing me, knowing you: Emotion differentiation in oneself is associated with recognition of others' emotions. *Cognition and Emotion*, *33*, 1461–1471. doi:10.1080/02699931.2019.1577221

Kang, S.-M., & Shaver, P. R. (2004). Individual differences in emotional complexity: Their psychological implications. *Journal of Personality*, *72*, 687–726.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(ECVP Abstract Suppl.), 14.

Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in Psychology*, *7*, Article 1324. doi:10.3389/fpsyg.2016.01324

Marsh, A., Ambady, N., & Kleck, R. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion*, *5*, 119–124.

Mellentin, A. I., Dervisevic, A., Stenager, E., Pilegaard, M., & Kirk, U. (2015). Seeing enemies? A systematic review of anger bias in the perception of facial expressions among anger-prone and aggressive populations. *Aggression and Violent Behavior*, *25*, 373–383. doi:10.1016/j.avb.2015.09.001

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Penton-Voak, I. S., Thomas, J., Gage, S. H., McMurran, M., McDonald, S., & Munafò, M. R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychological Science*, *24*, 688–697. doi:10.1177/0956797612459657

Poulin, F., & Boivin, M. (1999). Proactive and reactive aggression and boys' friendship quality in mainstream classrooms. *Journal of Emotional and Behavioral Disorders*, *7*, 168–177. doi:10.1177/106342669900700305

Price, R. B., Brown, V., & Siegle, G. J. (2019). Computational modeling applied to the dot-probe task yields improved reliability and mechanistic insights. *Biological Psychiatry*, *85*, 606–612. doi:10.1016/j.biopsych.2018.09.022

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

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922. doi:10.1162/neco.2008.12-06-420

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*, 260–281. doi:10.1016/j.tics.2016.01.007

Schönenberg, M., & Jusyte, A. (2014). Investigation of the hostile attribution bias toward ambiguous facial cues in antisocial violent offenders. *European Archives of Psychiatry and Clinical Neuroscience*, *264*, 61–69. doi:10.1007/s00406-013-0440-1

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

Tremblay, P. F., & Ewart, L. A. (2005). The Buss and Perry Aggression Questionnaire and its relations to values, the Big Five, provoking hypothetical situations, alcohol consumption patterns, and alcohol expectancies. *Personality and Individual Differences*, *38*, 337–346. doi:10.1016/j.paid.2004.04.012

Verona, E., & Bozzay, M. L. (2017). Biobehavioral approaches to aggression implicate perceived threat and insufficient sleep: Clinical relevance and policy implications. *Policy Insights from the Behavioral and Brain Sciences*, *4*, 178–185. doi:10.1177/2372732217719910

Vieira, J. B., Tavares, T. P., Marsh, A. A., & Mitchell, D. G. V. (2017). Emotion and personal space: Neural correlates of approach-avoidance tendencies to different facial expressions as a function of coldhearted psychopathic traits. *Human Brain Mapping*, *38*, 1492–1506. doi:10.1002/hbm.23467

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

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

Wilkowski, B. M., & Robinson, M. D. (2012). When aggressive individuals see the world more accurately: The case of perceptual sensitivity to subtle facial expressions of anger. *Personality and Social Psychology Bulletin*, *38*, 540–553.

Zimmer-Gembeck, M. J., & Nesdale, D. (2013). Anxious and angry rejection sensitivity, social withdrawal, and retribution in high and low ambiguous situations. *Journal of Personality*, *81*, 29–38. doi:10.1111/j.1467-6494.2012.00792.x