

YOUTHS' FACIAL APPEARANCE DISTINGUISHES LEADERS FROM FOLLOWERS IN GROUP-PERPETRATED CRIMINAL OFFENSES AND IS ASSOCIATED WITH SENTENCING OUTCOMES

JESSICA E. SUTHERLAND 

York University

ANA M. COJOCARIU

DAVID M. DAY

Ryerson University

ERIC HEHMAN

McGill University

Group-perpetrated crime often involves leaders and followers, but it is not currently understood how peer groups form around leaders during a criminal incident. Impression formation research has shown that specific facial cues are associated with leadership and perceptions of leadership. We extend this research to leadership among group-perpetrated youth crime and examine its role in downstream sentencing outcomes. Study 1 revealed that leaders of groups may be perceived as more dominant than their followers. In Study 2, participants were tasked with selecting the leaders from their groups and were more likely to (correctly) select targets perceived as more dominant but also (incorrectly) select targets perceived as more Trustworthy. In Study 3, we examined whether facial impressions were associated with downstream sentencing outcomes. Perceptions of Trustworthiness were associated with reduced sentencing, but dominance was unrelated. The results underscore the role that facial appearance plays in group formation and sentencing among youth.

Keywords: youth offending; criminal groups; impression formation; youth sentencing

Adolescence is a period of development marked by a significant increase in risk-taking behavior, which may include delinquency. Moffitt's (1993) dual taxonomy of offending proposes that adolescence-limited offenders, who comprise the vast majority of young people involved in some antisocial and delinquent behavior, begin offending during adolescence, cease offending in late adolescence, and engage in low-level offending alongside peers. By contrast, life course-persistent offenders, who comprise about 5% of adolescents involved in

AUTHORS' NOTE: *This research was supported by a SSHRC Insight Development Grant (430-2016-00094) to EH. Correspondence concerning this article should be addressed to Jessica E. Sutherland, Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3; e-mail: jesuther@yorku.ca.*

CRIMINAL JUSTICE AND BEHAVIOR, 201X, Vol. XX, No. X, Month 2019, 1–21.

DOI: 10.1177/0093854819889645

Article reuse guidelines: sagepub.com/journals-permissions

© 2019 International Association for Correctional and Forensic Psychology

delinquency and antisocial behavior, engage in more persistent delinquency throughout adolescence and into adulthood, and this pattern emerges at an earlier age. These dual pathways are supported by the age-crime curve, whereby criminal behavior generally declines after adolescence (Stolzenberg & D'Alessio, 2008).

A key distinction of crimes committed by adolescents is that they are more likely to occur alongside a peer group than are crimes by adults (Carrington, 2009; Moffitt, 1993). Compared with childhood and adulthood, adolescence is marked by a significant increase in susceptibility to peer influence (Albert et al., 2013), as well as emerging autonomy from parents and an increase in time spent with peers (Brown, 2004). These changes are underlaid by neurobiological evidence indicating that the onset of puberty is associated with significant reorientation of the adolescent brain toward social and emotional stimuli in their environments, such as indicators of social acceptance in a peer group (Forbes & Dahl, 2010). As a result, sensitivity to peer group dynamics, such as peer group norms and one's relative social status in a group, becomes heightened in adolescence (LaFontana & Cillessen, 2010; Shi & Xie, 2012). Thus, social status in a peer group is especially relevant for considering how leaders and followers may emerge, particularly in delinquent peer groups.

The significance of being a leader in a peer group engaged in delinquency has been highlighted by early work on cooffending that suggests some youth serve as recruiters to such a peer group while others are primarily joiners (Reiss, 1988). In contrast, Warr (1996) argues that while some youth are more likely to instigate crimes (e.g., because of more experience engaging in criminal behavior), leadership in cooffending contexts is primarily driven by situational and not stable individual traits. Some evidence exists for both possibilities. Criminal behavior in a group of youth will typically involve a leader and his or her followers (McGloin & Nguyen, 2012). Yet, group members who are older (Amemiya et al., 2016; van Mastrigt & Farrington, 2011; Warr, 1996) and who have more specific expertise related to criminality (i.e., having engaged in criminal behavior before; Amemiya et al., 2016) are more likely to be instigators. Behaviors specific to the incident can also be indicators of leadership. Instigators in gang rape cases (Porter & Alison, 2001) and group robbery (Porter & Alison, 2006) were identifiable by their level of involvement in the act and by their particular actions, including initiating the idea, selecting the victim, and disposing of the body when the crime resulted in death. Together, these findings suggest two trends: first, leaders do typically emerge in a group; and second, leaders can be identified by individual-specific factors (e.g., age and criminal expertise), and/or by the particular actions taken during the crime itself.

Although leadership during a criminal incident may be incident-specific or driven by individual characteristics, it is not well understood how peer groups coalesce around an instigator when initiating a particular crime. First, the effect of others' presence during a criminal act may reflect social facilitation or coercion effects, whereby group members' performances on a given task are enhanced by the presence of others (Geen, 1991). In the case of group-perpetrated crime in youth, leaders may instigate a crime due to the support of other group members, and followers may participate because other group members are also participating. Second, youth might select peers who are similar to their own level of delinquency (i.e., by seeking peers who engage in similar behaviors; Haynie et al., 2014) or they might be socialized into delinquency by spending time around peers engaged in delinquent behavior and adopting the delinquency-promoting norms of the group (Akers et al., 1979; Haynie & Osgood, 2005). For example, offending with others instead of

alone resulted in larger earnings from income-based crimes (Rowan et al., 2018), suggesting that selecting others who are competent at offending may increase the potential benefits of such activity. In contrast, McGloin and Thomas (2016) proposed that the benefits of offending must overcome the potential costs of offending for young people to consider engaging in crimes with others. In a group setting, the presence of others may reduce concerns about adverse outcomes and increase the salience of rewards (Rockloff & Dyer, 2007). The salience of social rewards during adolescence may therefore decrease concerns about sanctions and increase the likelihood of cooffending (da Silva et al., 2018; McGloin & Thomas, 2016) and subsequent socialization into peer group norms that support criminal behavior.

The selection-socialization pathways to delinquency appear to be complementary rather than distinct (Malloy et al., 2007; Monahan et al., 2009) and interaction between selection and socialization processes appear to better account for the dynamic nature of peer groups engaged in delinquency (Thornberry, 1987). For example, Malloy et al. (2007) found that youth in secure custody tended to fraternize with peers who had similar levels of delinquency as their own. However, youth also self-reported a greater variety of delinquent acts (e.g., violence and stealing) during their incarceration period, relative to their preincarceration period, when they had made new friends who also reported more variety in their preincarceration delinquency. Selection and socialization into criminal behavior are also sensitive to developmental stages: younger adolescents tended to form groups based on selection and socialization processes, whereas older adolescents and young adults were predominantly socialized into criminal behavior (Monahan et al., 2009).

Once a group has formed, however, it is not clear what factors influence how groups are structured during the commission of a particular criminal incident or what interpersonal processes may underlie group members' willingness to join. Evidence from noncriminal contexts indicates that not all youth are similarly able to influence others to engage in particular behaviors (Allen et al., 2012; Choukas-Bradley et al., 2015) and that social status and popularity in a peer group are predictors of a group member's ability to influence aggression in other group member (Shi & Xie, 2012). Youth convicted of sexual offenses committed with a group reported feeling social pressures, fear of social rejection, or direct orders from group members to engage in the act (da Silva et al., 2018), suggesting that incident-specific interpersonal dynamics can be powerful influences on criminal behavior beyond broader peer selection and socialization processes. Thus, consistent with the threshold model proposed by McGloin and Rowan (2015) and Thornberry's (1987), interactional perspective, individual, and situational characteristics can independently and interactively increase the likelihood of cooffending overall, but also may determine group structure in a given criminal incident.

It is possible that psychological processes, such as impression formation, may interact with situational factors and influence the likelihood of a group member becoming a leader in a criminal incident while others are relegated to follower roles. Research in social and cognitive psychology has explored how groups form and operate and, particularly, how groups coalesce around a leader. Facial characteristics, such as attractiveness and dominance, have been found to distinguish group leaders from followers (Olivola, Eubanks, & Lovelace, 2014; Re & Rule, 2017). Insights yielded by this work may inform research on youth group-perpetrated crime but have generally not been brought to bear in understanding group formation in adolescent offending. In an individual criminal incident, it may be that

group members provide or accept direction based on processes that help group members form impressions of others' competence or dominance in the given incident.

THE ROLE OF PERCEIVED FACIAL APPEARANCE IN GROUP DYNAMICS

Recent research in social psychology has suggested that physical appearance plays a role in the process of group formation. Perceivers use physical appearance, particularly of the face, to infer the characteristics of others (for review, see Todorov et al., 2015). For example, individuals with more prominent brows or wider faces relative to its height are perceived as more dominant than individuals with less prominent brows and narrower faces (Hehman et al., 2013). Importantly, although these inferences into character are not always accurate (Olivola & Todorov, 2010b), there is often strong consensus among perceivers about the features that correspond to personality characteristics (Todorov et al., 2015). Accordingly, because facial appearance is used to infer character, individuals use these physical features to guide decisions when forming groups (Hehman et al., 2015; Olivola, Funk, & Todorov, 2014) and accordingly, individuals within existing groups physically resemble one another (Hehman et al., 2018).

Despite within-group similarity, there is also within-group variability in group members' appearances. This variability is important, because distinct social roles within a group are associated with meaningful differences in appearance. Specifically, group leaders tend to have facial characteristics that differ from other group members. In previous research, group leadership was associated with facial features eliciting perceptions of dominance and competence (Re et al., 2013; Re & Rule, 2017), such as the overall length of the face (Re et al., 2013), and maturity (Cherulnik et al., 1990). Furthermore, individuals perceived to be more competent and dominant were more likely to be selected for leadership in a political context (Olivola & Todorov, 2010a; Todorov et al., 2005; Zebrowitz & Montepare, 2005) and to be chief executive officers (Graham et al., 2017). In general, this research reveals that group leaders in some domains can be distinguished from other group members by their facial characteristics (e.g., Re & Rule, 2017; Van Vugt & Grabo, 2015).

Relative to other group members, leaders appear to have unique facial characteristics that elicit perceptions of traits considered advantageous for leadership (e.g., dominance). However, which specific characteristics are considered desirable depends on the context of the group and the leadership role. For instance, participants preferred leaders who evoked perceptions of dominance and masculinity during wartime, but preferred leaders who evoked perceptions of trustworthiness during peacetime (Grabo & Van Vugt, 2018; Little et al., 2007; Van Vugt & Grabo, 2015). Similarly, in conflict-anticipatory contexts, individuals wanted more dominant and strong-looking individuals to join the group (Hehman et al., 2015). Individuals were able to distinguish between leaders and nonleaders across multiple domains (e.g., sports, business, and the military), but could also distinguish leaders from followers within each specific domain, supporting the notion that leaders in different contexts have facial features that elicit distinct impressions (Olivola et al., 2014; Re & Rule, 2017).

THE ROLE OF PERCEIVED FACIAL APPEARANCE AMONG YOUTH CRIMINAL GROUP MEMBERS

If criminal incidents can be characterized as conflict-anticipatory contexts, leaders may demonstrate similar characteristics to leaders in noncriminal contexts, such as dominance

and strength, relative to their group members. In other words, impression formation processes may also factor into groups forming for the purpose of criminal behavior and coalescing around a leader. For example, a leader's physical appearance may convey strength or experience, attracting youth to join the leader's peer group (e.g., Amemiya et al., 2016; McGloin & Nguyen, 2012) and increasing the likelihood of criminal behavior through the socialization pathway (Malloy et al., 2007; McGloin & Rowan, 2015; Monahan et al., 2009). It may also be that dominant leaders attract group members who are more submissive, supporting the concept of interpersonal complementarity (i.e., dominant and submissive interpersonal styles) as predictive of successful group formation (Dryer & Horowitz, 1997). McGloin and Rowan (2015) have also proposed that being part of a group may prime preferences for ingroup cohesion, possibly increasing the likelihood that members will cooffend to achieve social rewards (McGloin & Thomas, 2016).

Furthermore, the facial appearance of group members involved in a group-perpetrated criminal incident may influence sentencing outcomes in the judicial system. Previous research among adult populations has revealed that facial cues provide an extralegal influence on criminal justice outcomes. For example, in small-claims court, individuals with baby-faced features (e.g., large eyes, round cheeks), eliciting impressions of trustworthiness and naïveté, were more likely to win their lawsuits (Zebrowitz & McDonald, 1991) compared with individuals with mature facial features (e.g., square jaws, angular faces). Similarly, Flowe (2012) found that impressions of trustworthiness were negatively associated with perceived criminality in a sample of adult offenders. Moreover, Wilson and Rule (2015) found that being perceived as less trustworthy was predictive of a greater likelihood of receiving a death sentence relative to a life sentence. Perceptions of trustworthiness correlate highly with attractiveness (Oosterhof & Todorov, 2008), and attractiveness has a similar buffering effect on sentence severity, with more attractive defendants receiving less severe sentences (Stewart, 1980). No study has yet examined the influence of appearance on sentencing outcomes among youth in conflict with the law, yet similar results might be expected. Consistent with this idea, attorneys report that they believe youth will be viewed as less culpable when they have a more youthful, rather than mature, appearance (Camilletti & Scullin, 2012). In the context of cooffending, leaders perceived as less dominant or competent may subsequently receive a lighter sentence, despite their culpability in the crime relative to other group members. Addressing this important gap in the literature, the present work examined the role of appearance on sentencing outcomes among a sample of youth involved in the criminal justice system.

THE PRESENT RESEARCH

Determining whether leaders differ in appearance from followers in group-perpetrated crime is important because of the potential influence of physical appearance on two domains relevant to understanding groups in criminal contexts. First, differences in appearance may account for how these groups might initially form and commit crimes together, and second, such differences may account for how group members are treated by the justice system. Therefore, we addressed three questions within the context of youth group-perpetrated crime: (a) whether leaders of groups were distinguishable from followers by their facial appearance, (b) whether perceivers were able to use these potential facial cues to accurately distinguish leaders from followers, and (c) whether these facial

cues were associated with sentencing outcomes in judicial proceedings. These questions are examined in three studies.

STUDY 1

Study 1 tested whether leaders differed from followers in facial appearance as determined by trait ratings.

METHOD

Stimuli

Facial stimuli were obtained through a systematic search of all adolescent and young adult group crime cases reported in U.S. daily newspapers in all 50 states. Newspaper coverage of youth crime often focuses on specific episodes of youth crime rather than systemic concerns about it (Ruigrok et al., 2017) and therefore typically emphasizes the severity of and presence of violence in youth crime (Faucher, 2009). Thus, daily newspapers were an ideal source for locating a representative sample of cases of serious youth crime and relevant information, such as group formation and relationships among group members.

Daily newspapers were accessed through ProQuest Newsstream and corresponding newspaper websites. Search terms used were as follows: teens, youths, arrested, and charged. Date limits were set to anything after January 1, 2000, until November, 2017, to maximize the availability of newspapers' coverage of a case as digital news coverage became more common. We excluded articles about single youth or youths above the age of 25 due to evidence indicating that peer influence on delinquency extends into early adulthood (e.g., Monahan et al., 2009), and/or articles that did not mention the names or mug shots/photos of youth involved in the alleged crimes. Our search strategy identified 1,192 news articles that met our basic inclusion criteria of involving at least two youths below the age of 25, including the names of all involved, and/or including mug shots of group members.

All articles that met inclusion criteria were coded for newspaper title, year of publication, number of youth, age of youth, gender of youth, and criminal offense(s). Once the search was complete, articles were reviewed and coded by two reviewers for availability of photos for all youth within each group involved. If mug shots/photos of the youth were included in the article, they were kept; if only names were provided, state correctional databases were searched to locate mug shots or booking photos. If no photos could be located despite availability of names, the case was excluded. A total of 213 individuals in 75 groups (i.e., individual criminal cases) were identified. Of all photos obtained, 205 (96.2%) were mug shots, four (1.9%) were of the group member in court, and four (1.9%) were taken elsewhere (e.g., a personal photograph). Photos were saved and cropped to the face. All stimuli are available here [https://osf.io/b98cm/?view_only=9624dc2af04d43f086f8f750cb207aa].

Groups varied in size, ranging from two to seven members ($M = 2.84$, $SD = 1.15$) and 91% of the youth were male. Sixteen groups (21.3%) were mixed gender. Individuals were coded as White or non-White based on visual examination of each photo by one author, and 23 groups (30.1%) had multiple races represented. With the exception of five cases that involved nonviolent index offenses (one theft, one series of vehicle burglaries, and three cases of intentional arson), all index offenses in each case involved serious violent crimes

against victims, including aggravated robbery, assault, sexual assault, and murder. This is consistent with Ruigrok et al. (2017), who noted that media coverage of youth crime tends to emphasize serious and violent incidents.

Determining Leadership Status

For each case, the group leader was identified by the first two authors from facts about the case contained within the news stories. Specifically, we classified a group member as “leader” if an investigating officer, prosecutor, judge, or other court actor was on the record directly naming or reporting suspicion that group member as the ringleader of the crime(s) in question. Group members who were named as leaders had typically selected the victim and/or initiated the crime in question, instructed other group members to take particular roles in the crime, and/or determined how to leave the scene and victim, consistent with definitions of criminal leadership by Porter and Alison (2001, 2006). This method of leadership selection was used to determine how “leadership,” as defined by criminal justice actors, affected perceptions of faces and downstream court actions. Identifying group leaders was done prior to faces being rated and any analyses conducted.

Participants and Procedure

A total of 1,014 participants were recruited from Amazon Mechanical Turk for monetary compensation. Participants completed a 5-min task in which they rated faces on one trait (attractiveness, competence, dominance, physical strength, trustworthiness, or warmth). These traits generally capture the dimensions thought to comprise the facial impression space (Oosterhof & Todorov, 2008), and some have previously been found to be associated with leadership (Olivola et al., 2014; Re et al., 2013; Re & Rule, 2017). Participants rated faces on only one of the six traits in a between-subject design, to avoid previous ratings biasing subsequent ratings. Participants saw individual mug shots, randomized by participant, and rated them on one of the six traits using a scale of 1 (“Not at all”) to 7 (“Very much”), prior to completing a brief demographics questionnaire. To limit participant fatigue, each participant rated 71 photos. Therefore, each of the 213 mug shots was rated on a single trait by at least 40 different participants.

Data were cleaned according to our preregistered data-cleaning procedure [osf.io/65tpb/]. Participants were removed from analysis if they did not finish the task, had more than 20% repeated responses (e.g., rating subsequent faces a “1” repeatedly), or had more than 20% of ratings completed in less than 400 ms, leaving 943 participants for analysis (attractiveness, $n = 151$; competence, $n = 156$; dominance, $n = 177$; strength, $n = 165$; trustworthiness, $n = 157$; warmth, $n = 137$) and 57,398 individual ratings across the six traits. Participants in this final sample ($M_{\text{age}} = 36.16$, $SD = 12.62$) were primarily female (60%) and White (73.2%). Ratings were averaged across participants to create average trait ratings for each target face, which functioned as the unit of analysis.

Creating Dimensions

We anticipated that ratings for some of the traits would be correlated (e.g., “strong” and “dominant”). Traditionally, researchers studying this impression formation space collect ratings of targets on various traits and use data-reduction techniques, like factor analysis,

to create the dimensions underlying impression formation. Most research finds two primary dimensions comprising this impression formation space (Oosterhof & Todorov, 2008; Sutherland et al., 2013). One dimension, typically labeled Trustworthiness, reflects whether a target's intentions appear to be positive or negative toward the perceiver. Traits such as trustworthiness, warmth, and attractiveness load on this dimension. A second dimension, often labeled Dominance, is thought to reflect a target's ability to enact those intentions. Traits like dominant or physically strong load on this (Oosterhof & Todorov, 2008; Sutherland et al., 2013).

As expected, bias-corrected bootstrapped correlations revealed that perceptions of dominance and physical strength were highly correlated ($r = .741$, 95% confidence interval [CI] = [.679, .792], $p < .001$), as were perceptions of warmth, competence, trustworthiness, and attractiveness (all r s $> .551$, all p s $< .001$; Table 1). Principal axis factoring with oblimin rotation suggested a two-factor underlying structure based on scree plot and eigenvalues. Accordingly, dimensions were created by averaging ratings of physical strength and dominance, and averaging ratings of warmth, competence, attractiveness, and trustworthiness. These dimensions are henceforth referred to as Dominance and Trustworthiness, respectively, and were used in subsequent analyses to examine traits associated with leadership status and criminal sentencing in the youth criminal groups.

Analytic Approach

Data were analyzed in a multilevel framework, as individual targets were nested within groups. Groups, therefore, served as the clustering variable in Study 1. Intercepts were random and slopes were fixed. Targets varied by gender, age, and racial majority/minority status within groups, and groups themselves varied in size. We included these variables in the models as covariates throughout all analyses and all variables and covariates were entered simultaneously in the models, though excluding these covariates does not change conclusions. Racial majority status was defined as White, and minority as all other races/ethnicities. Dichotomous variables were represented with contrast codes (female = -1, male = 1; racial minority = -1, racial majority = 1). Level 1 variables (i.e., Dominance and Trustworthiness scores) were group-mean centered, whereas Level 2 variables (i.e., group size) were grand-mean centered, recommended as best practice (McCoach & Adelson, 2010). Analyses were completed in R using lme4 (Bates et al., 2014). The data are available here [osf.io/b98cm/].

RESULTS

To test whether leaders could be differentiated from their followers by their facial features, we regressed whether targets were leaders or followers on the Dominance and Trustworthiness dimensions alongside covariates in a multilevel logistic model. Results indicated that leaders were not perceived as significantly more Dominant than followers ($B = .56$, $SE = .31$, $p = .076$; Figure 1; Table 2), and older group members were also not significantly more likely to be leaders ($B = 0.27$, $SE = 0.15$, $p = .070$; Figure 1; Table 2) at the standard alpha level of .05. Ratings of Trustworthiness were unrelated to leadership/follower status.

Groups were mixed race and gender, and we repeated the same analysis on a smaller data set restricted to groups of White males only, consisting of 53 targets across 21 groups (Table 3). Similarly, leaders were not rated significantly more Dominant than followers ($B = 1.08$,

TABLE 1: Correlation Matrix of Primary Variables Involved in Studies 1 and 3

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Follower/Leader	—												
2. Group size	-.25**	—											
3. Target gender	-.02	.00	—										
4. Target race	-.00	.12	-.05	—									
5. Attractiveness	-.02	.12	.32**	.10	—								
6. Dominance	.07	.15*	-.23**	.29**	-.11	—							
7. Physical strength	.07	.16*	-.45**	.39**	-.03	.74**	—						
8. Trustworthiness	-.04	-.04	.11	.14*	.60**	-.40**	-.12	—					
9. Warmth	-.06	.02	.14*	.14*	.55**	-.52**	-.16*	.82**	—				
10. Competence	.02	.06	.13	.01	.61**	-.20**	-.09	.81**	.68**	—			
11. Dom/Strong	.07	.17*	-.38**	.37**	-.07	.92**	.94**	-.27**	-.35**	-.15*	—		
12. Attract/Trust/ Warm/Competence	-.03	.05	.20**	.12	.80**	-.36**	-.12	.93**	.88**	.88**	-.24**	—	
13. Age	.07	.12	-.02	.19**	-.06	.17*	.16*	-.08	-.05	-.00	.17*	-.06	—

* $p < .05$. ** $p < .01$.

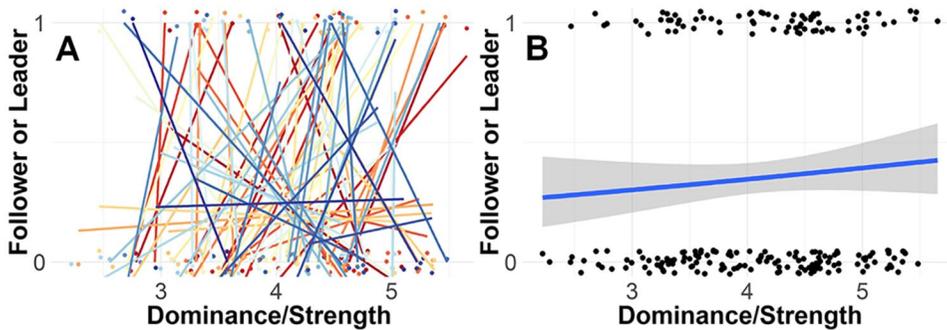


Figure 1: Relationship Between Ratings of Dominance/Physical Strength and Actual Leadership

Note. Panel A spaghetti plot demonstrating the variability of this relationship across criminal youth groups. Each specific group is represented by same-color lines and points. Panel B shows the average trend across all groups, with the shaded area representing a 95% confidence band of the estimate.

TABLE 2: Predicting True Leadership Status

True leadership status	<i>B</i>	<i>SE</i>	<i>p</i>
Dominance/Strength	0.56	0.32	.077
Age	0.27	0.15	.076
Trust/Attract/Warm/Competence	0.06	0.45	.894
Target majority/minority Status	-0.02	0.16	.876
Target gender	0.09	0.29	.761
Group size	-0.45	0.13	<.001

$SE = 0.56, p = .054$) at the standard alpha level of .05, and age was again unrelated ($B = 0.19, SE = 0.36, p = .593$). Trustworthiness was also unrelated ($B = -0.98, SE = 0.80, p = .220$) to leader/follower status.

TABLE 3: Predicting True Leadership Status, in Groups of White Men Only

True leadership status for groups of White men	<i>B</i>	<i>SE</i>	<i>p</i>
Dominance/Strength	1.08	0.56	.054
Trust/Attract/Warm/Competence	-0.98	0.80	.220
Age	0.19	0.36	.593
Group size	-0.36	0.28	.195

DISCUSSION

In summary, in Study 1 we find leaders do not look significantly different from followers on Dominance or Trustworthiness dimensions at the standard alpha level of 0.05; however, as seen in Figure 1, leaders were generally perceived as more dominant than followers. This is in line with previous research indicating that leaders are distinguishable from other group members based on physical characteristics denoting strength and dominance (e.g., Grabo & Van Vugt, 2018; Little et al., 2007; Re & Rule, 2017; Van Vugt & Grabo, 2015). In addition, despite no significant effect of age predicting leadership at the standard alpha level of .05, as previously found in Amemiya et al. (2016), van Mastrigt and Farrington (2011), and (Warr, 1996), leaders may still be more likely to be older than followers. Leaders are more likely to have criminal expertise (Amemiya et al., 2016; Rowan et al., 2018), and it is likely that age and criminal expertise positively correlate and leaders with more criminal expertise may be more likely to attract group members.

STUDY 2

While Study 1 examined potential appearance-based difference in leaders and followers, Study 2 examined what appearance-based cues perceivers might use when assessing leadership. Thus, whereas Study 1 focused on distinguishing actual leaders from followers, Study 2 focused on the determinants of perceptions of leadership when perceivers are directly asked to select a leader.

METHOD

Participants and Procedure

Two hundred and fifty-four U.S. participants were recruited from Amazon Mechanical Turk to complete a 10-min task in exchange for monetary compensation. Participants were told that they would be asked to select a leader from a group of mug shots. Participants were presented with arrays of the same faces depicting each of the 75 groups from Study 1 and selected one face from each group who they thought was the leader. Therefore, the number of faces presented in each array varied with the size of the group. They subsequently completed a brief demographics questionnaire. The left-to-right order of the photographs displayed within a group was randomized by participant, as was the order of the groups themselves. Data were cleaned consistent with our preregistered cleaning method, which left 204 participants for analysis. Participants ($M_{\text{age}} = 41.27$, $SD = 14.13$) were primarily female (66.54%) and White (77.56%).

TABLE 4: Predicting Perceived Leadership Status Based on Facial Characteristics

Perceived leadership status	<i>B</i>	<i>SE</i>	<i>p</i>
Dominance/Strength	1.48	0.11	<.001
Trust/Attract/Warm/Competence	1.23	0.14	<.001
Target majority/minority status	-0.39	0.02	<.001
Target gender	0.20	0.03	<.001
Group size	-0.53	0.04	<.001
Age	0.12	0.01	<.001
Dominance × Trust	-0.15	0.03	<.001

Analytic Approach

We combined the averaged ratings of Dominance and Trustworthiness from Study 1 with the leadership decisions made by participants in Study 2 to examine which trait dimensions (measured in Study 1) were driving leader selection in Study 2. To test which factors were driving leadership decisions, data were analyzed in cross-classified multilevel models (Kenny, 1991; Raudenbush & Bryk, 2002). Because leadership decisions were repeatedly made by participants and were made repeatedly of targets within groups, data were nested both within participants and youth criminal groups. Analyzing the data in this manner yields appropriate standard errors due to this clustering and allows conclusions from this research to generalize beyond the specific participants and stimuli (Judd et al., 2012). Leadership decisions (0 = Not selected as leader, 1 = Selected as leader) were regressed on Dominance and Trustworthiness dimensions entered simultaneously alongside covariates in a logistic model. Intercepts were random and slopes were fixed. The data are available here [osf.io/b98cm/].

RESULTS

Because there were 75 groups and 213 members, chance selection of leadership across all groups was ~35%. On average, participants were below chance, classifying leaders correctly on only 16% of trials. Turning to our main hypotheses, we examined which traits led to leadership selection (Table 4). Results revealed that faces rated higher on Dominance were more likely to be selected as leader ($B = 1.48$, $SE = 0.11$, $p < .001$). In addition, faces rated higher on Trustworthiness were more likely to be selected as leader ($B = 1.22$, $SE = 0.13$, $p < .001$), although the effect of Dominance was larger. These main effects were qualified by a Dominance × Trustworthiness interaction ($B = -0.15$, $SE = 0.03$, $p < .001$). We interpreted this interaction at $\pm 1 SD$ of the Trustworthiness dimension. Simple effects revealed that when a face was 1 SD below average on Trustworthiness, perceptions of Dominance were used as a slightly stronger indicator of leadership ($B = 1.49$, $SE = 0.14$, $p < .001$) than when a face was 1 SD above average on Trustworthiness ($B = 1.36$, $SE = 0.10$, $p < .001$). We detected this interaction due to our high statistical power, though generally perceptions along the Trustworthiness dimension did not dramatically shift the relationship between Dominance and leadership choices.

Like in Study 1, we replicated our findings on a smaller data set of groups comprised of White males only (Table 5). We found the same effects in this smaller data set, such that Dominance ($B = 2.98$, $SE = 0.24$, $p < .001$) and Trustworthiness ($B = 2.60$, $SE = 0.26$,

TABLE 5: Predicting Perceived Leadership Status Based on Facial Characteristics in Groups of White Men Only

Perceived leadership status for groups of white men	<i>B</i>	<i>SE</i>	<i>p</i>
Dominance/Strength	2.98	0.24	<.001
Trust/Attract/Warm/Competence	2.60	0.26	<.001
Group size	-0.60	0.10	<.001
Age	-0.04	0.03	.155
Dominance × Trust	-0.50	0.07	<.001

$p < .001$) had significant main effects on perception of leaders and similarly interacted ($B = -0.50$, $SE = 0.07$, $p < .001$).

DISCUSSION

Study 2 indicates that Dominance and Trustworthiness were each independently associated with lay perceptions of leadership. Importantly, this pattern of results diverges from Study 1, in which leaders were not significantly more Trustworthy- or Dominant-looking. Importantly, perceptions of leadership and leaders' appearances might influence decisions of court actors. Accordingly, Study 3 examined how perceptions of leadership and appearance were associated with courtroom outcomes.

STUDY 3

In the context of the adult criminal justice system, appearance is associated with sentencing outcomes (Stewart, 1980; Zebrowitz & McDonald, 1991). Testing whether this possibility extends to youth convicted of a group-perpetrated crime, we examined whether perceptions of Dominance and Trustworthiness were associated with sentencing decisions.

METHOD

Procedure

Sentencing information was collected from the original news sources and from follow-up articles that reported on case outcomes of each group and group member. Follow-up articles were sought by specifically using Google's search engine to find additional information about each case. Each group member's name was searched alongside a "sentencing" qualifier to identify follow-up information. We also searched for sentencing information in states which make inmate information available in corrections databases if it was not available in news articles, though this strategy was limited only to adults' sentences and only to certain states. Court and sentencing outcomes were coded for the type of sentence (e.g., custody, probation) and length of sentence in years by one research assistant uninvolved in any prior data collection, and blind to the results of Studies 1 and 2. Targets were the same as in Studies 1 and 2, but a reduced sample of 121 mug shots was used in Study 3, due to availability of sentencing and court outcome information. For some groups, sentencing information was not reported, whereas for others, the case was still ongoing for various group members and so complete information was not available. Complete sentencing information was available for 121 (of 213) group members across 47 groups (of 75).

TABLE 6: Predicting Sentence Length Based on Facial Characteristics

Log-transformed sentence length	<i>B</i>	<i>SE</i>	<i>p</i>
Dominance/Strength	0.07	0.14	.596
Trust/Attract/Warm/Competence	-0.43	0.20	.031
Leader/Follower	0.29	0.07	<.001
Target majority/minority status	0.03	0.10	.686
Target gender	-0.15	0.33	.657
Group size	-0.05	0.10	.577
Age	-0.04	0.06	.495

Analytic Approach

Because we were focusing on sentencing outcomes associated with group-perpetrated crime, we returned to the multilevel analytic approach adopted in Study 1. Sentencing outcomes were examined in a multilevel model with group members nested within their groups. This approach takes into account that youth within a group engaged in crime together, and thus their sentencing decisions are not independent. Again, continuous Level 1 variables were group-mean centered and Level 2 variables were grand-mean centered. Whether targets were actually followers or leaders, was additionally included as a contrast code (followers = -1, leaders = 1) and all independent variables and covariates were entered simultaneously. Intercepts were random and slopes were fixed.

Lengths of sentences varied dramatically across group members, ranging from 1.5 years to life without parole to death. For purposes of analysis, we converted these sentences to a numeric value representing years in custody. Some sentences were ranges (e.g., 18–45 years). The range indicates that the individual will receive at least 18 years of prison time, but options for release may then be available until the final 45 years had been served. For these cases, we operationalized the sentence as the average of these years (e.g., 31.5 years). Other sentences did not include year estimates (e.g., life without parole). To represent that this sentencing outcome was considered a more severe penalty than the highest nonlife sentence of 75 years, we assigned all life ($n = 21$) and death ($n = 1$) sentences a value of 80 years. This variation in sentencing type reflects different trends between states in the use of life sentences and life without parole given to youth in conflict with the law (Rovner, 2017). However, our sample also contained young adults over the age of 18, making them eligible for sentences that may otherwise not be available to youth under the age of 18 as defined by the U.S. government. The resulting sentencing variable had a nonnormal distribution (skewness = .807). To meet statistical assumptions, we performed a log-transform and this version of the sentencing outcome variable was used as the outcome measure in analyses. The data are available here [osf.io/b98cm/].

RESULTS

Not surprisingly, results indicated that leaders received longer sentences than followers ($B = 0.29$, $SE = 0.07$, $p < .001$; Table 6). Relevant to our primary hypotheses, we found that facial appearance was associated with sentencing outcomes, though not in ways consistent with the perceptions of leaders and followers in Study 1 and 2. Specifically, perceptions of Dominance were unrelated to sentencing outcomes ($B = 0.07$, $SE = 0.14$, $p = .596$),

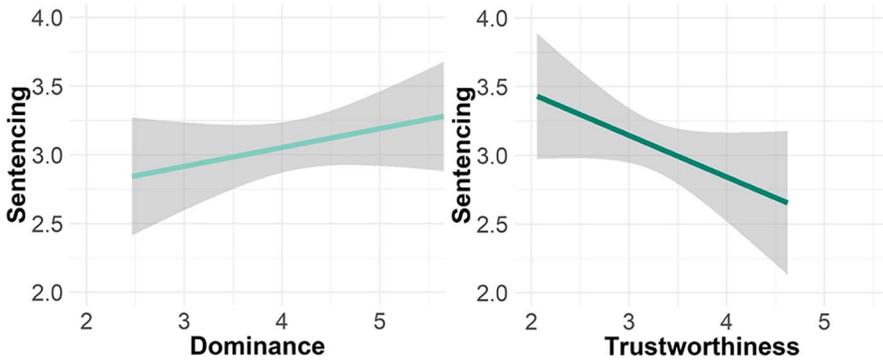


Figure 2: Relationships Between the Dominance and Trustworthiness Factors With Sentencing Outcomes

Note. The shaded area represents a 95% confidence band of the estimate.

TABLE 7: Predicting Sentence Length Based on Facial Characteristics in Groups of White Males Only

Log-transformed sentence length for groups of white men	<i>B</i>	<i>SE</i>	<i>p</i>
Dominance/Strength	0.06	0.26	.810
Trust/Attract/Warm/Competence	-0.90	0.34	.018
Leader/Follower	0.19	0.15	.230
Age	-0.15	0.14	.322
Group size	-0.24	0.27	.392

whereas perceptions of Trustworthiness were associated with reduced sentence lengths ($B = -0.43, SE = 0.19, p = .031$; Figure 2).

Again testing these findings on groups with only White male members, the results are largely consistent with those above (Table 7). Dominance did not predict sentence length ($B = 0.06, SE = 0.26, p = .810$), whereas Trustworthiness predicted shorter sentences ($B = -0.90, SE = 0.34, p = .018$). However, true leadership status ($B = 0.19, SE = 0.15, p = .230$) and age ($B = -0.15, SE = 0.14, p = .322$) did not predict longer sentences among groups of White males only.

DISCUSSION

In general, we show evidence that appearance introduces bias in courtroom decisions, but not in a way that is consistent with actual differences between leaders and followers. Namely, although it is possible that leaders are more Dominant in appearance than followers, there are no differences in perceived Trustworthiness (Study 1). Yet in the courtroom, perceived Dominance was unrelated to outcomes, while Trustworthiness was associated with reduced sentencing.

GENERAL DISCUSSION

The present studies extended current research in social cognition and impression formation to youth peer groups involved in criminal behavior for the first time, and examined whether

leaders in this context are perceived differently than followers. We also examined whether appearance-based differences were associated with downstream consequences of sentencing outcomes. Three key findings emerged. First, facial ratings along the Dominance dimension were not significantly associated with actual leadership status in group-perpetrated crime, though Figure 1 indicates leaders were generally perceived as more Dominant in appearance. This is consistent with previous literature about the facial traits of leaders in conflict-heavy contexts and indicates that similar patterns may also be present in crimes committed by groups of youth. Second, perceivers used both Dominance and Trustworthiness dimensions when selecting who they believed to be the leader among group members. Individuals higher on each dimension were more likely to be selected as the group leader, though Trustworthiness was unrelated to actual leadership status. Finally, having facial cues that give rise to greater perceptions of Trustworthiness was associated with reduced sentencing outcomes, while Dominance was unrelated. This is significant, because Trustworthiness was unrelated to true leadership status as found in Study 1 and may represent an extralegal factor influencing sentencing decisions.

Although it did not reach statistical significance (i.e., at the standard alpha level of .05), our finding that leaders within groups may be more Dominant in appearance than followers as seen in Figure 1 is consistent with previous literature indicating that group leaders are more likely to be rated as Dominant in legal and criminal domains (Re et al., 2013; Re & Rule, 2017). Furthermore, laboratory research in which participants choose group members indicates that Dominant group members are more likely to be selected as teammates in competitive, but not cooperative, contexts (Hegeman et al., 2015). What leadership in group-perpetrated youth crime, the mafia, law firms, competition, and countries during wartime have in common is that they are all contexts associated with the expectation of conflict. The evidence from across these multiple domains and studies indicates that individuals tend to coalesce around Dominant-looking leaders in conflict-heavy contexts.

This finding provides insight into how groups might form for the purpose of engaging in criminal behavior initially. In the current study, group leaders may appear more Dominant than followers, which may be interpreted in at least two ways. First, less Dominant-looking followers may coalesce around a more Dominant-looking leader, and groups do not form around less Dominant-looking leaders. Similar to research on political leader selection during wartime (Grabo & Van Vugt, 2018; Van Vugt & Grabo, 2015), individuals may feel that Dominant-looking individuals are more capable of success in conflict-oriented contexts and be more likely to follow these individuals. Consistent with the threshold hypothesis (McGloin & Rowan, 2015), it may be that Dominant-looking leaders inspire confidence in or the likelihood of benefiting from crime, increasing the likelihood that followers will coalesce around them. In addition, their Dominant appearance might increase the likelihood that group members submit to their directions or pressure during a criminal incident (da Silva et al., 2018).

Alternatively, it is possible that the groups included in the present research did not form with the intention to commit crime; rather they may have already existed as a peer group. That we found a possible effect of dominance on leadership that did not reach statistical significance may reflect this interpretation. Dominant individuals within these groups may have been effective at enticing and/or coercing other group members to engage in criminal acts and increasing their peers' susceptibility to their influence. By contrast, less Dominant leaders in other friendship groups may have been less successful in encouraging group

members to engage in crime; therefore, we may have observed a relationship between group leaders and Dominance. The difference in these two interpretations hinges on whether we assume the groups formed for the purpose of engaging in crime (i.e., the peer selection pathway), or whether more dominant leaders of already existing groups were only more successful at encouraging the group toward crime, but subsequent to the group's formation (i.e., the peer socialization pathway).

These interpretations are not exclusive. While the present data cannot speak to likelihood of these two possibilities, the former interpretation is consistent with previous research finding that individual characteristics of a group member, in this case their physical appearance, are associated with their likelihood of instigating a criminal incident (Amemiya et al., 2016; McGloin & Nguyen, 2012; van Mastrigt & Farrington, 2011). The latter interpretation, however, highlights the role that facial appearance and physical characteristics may contribute to motivating other group members and "tip the balance" toward offending (McGloin & Rowan, 2015; McGloin & Thomas, 2016), perhaps through increasing other group members' confidence or increasing compliance with group norms. Furthermore, the present study complements social development literature that finds that peer group influences on delinquent behavior may not only depend on selection and socialization processes (e.g., Monahan et al., 2009; Wong et al., 2013), but also on differential susceptibility to peer influence (e.g., Allen et al., 2006; Shi & Xie, 2012), including impression formation processes.

This outcome of selecting or coalescing around a leader as seen in Study 1 differed from participant decisions in Study 2, however. Participants in Study 2 were more likely to believe that leaders were more Dominant- and Trustworthy-looking than were followers. Thus, using a Brunswikian lens model conceptualization (Brunswik, 1956), perceptions of Dominance might be used by perceivers to accurately ascertain leadership within a group, but that perceptions of Trustworthiness might serve as a false signal, leading to inaccurate inferences. It is possible that when asked to select a leader, perceivers rely on other indicators of leadership in non-conflict-heavy contexts (e.g., attractiveness or trustworthiness).

Finally, perceptions of group members' facial appearance were associated with their sentencing outcomes. Study 3 revealed that group members who were rated higher on the Trustworthiness dimension received shorter sentences. In the United States, whereas juries decide whether defendants in a criminal case are guilty or not, it is judges who ultimately decide the severity of the punishment (with the exception of capital cases). Thus, our results suggest that judges may be swayed by defendants' appearance when making sentencing decisions, with individuals higher on the Trustworthiness dimension receiving lighter sentences. This result is consistent with other research showing that plaintiffs with more baby-faced features were more likely to win lawsuits filed against them due to perceptions of Trustworthiness being associated with involuntary negligence (i.e., that younger looking individuals were not purposely being negligent and were therefore less culpable; Zebrowitz & McDonald, 1991). Higher perceived Trustworthiness is also associated with a lower likelihood of receiving a death sentence (Wilson & Rule, 2015) and perceptions of criminality (Flowe, 2012). In the current study, it is important to note that Trustworthiness does not distinguish true leaders from followers within these groups, but is being used by perceivers (e.g., judges in criminal cases) when meting out punishment. This means that although true leaders may be

receiving sentences proportionate to their role in the crime relative to their group members, more Trustworthy-looking leaders are receiving lighter sentences, despite the fact that Trustworthiness is unrelated to leadership status and culpability. To the best of our knowledge, this is the first time that facial traits, such as Trustworthiness, have been empirically explored in the context of juvenile sentencing.

LIMITATIONS

The groups involved in the present research are not representative of all crimes committed in groups of youth, and the extent to which our conclusions generalize beyond these groups committing violent crimes is unclear. However, given other research finding preferences for dominant leaders in conflict-based contexts (Little et al., 2007), we speculate that our results would generalize to other delinquent groups in which conflict is anticipated. This conclusion may be tested by future research and compared across criminal domains, including violent, white-collar, property, and drug-related crime, in which different leadership traits may be perceived as more or less advantageous. For instance, in white-collar financial crime, perhaps intelligence or competence is a trait more strongly associated with leadership. Furthermore, it is possible that in gangs, in which leadership has been associated with more psychopathic traits and higher self-esteem (Dmitrieva et al., 2014), similar findings around competence and dominance would emerge. However, the organizational structure of gangs can occasionally be similar to the informal structure of peer groups of young people, where members join and leave over time (Weerman et al., 2015). Considering a single criminal incident and the group members involved as the unit of analysis offers important insights about the nature of situational sources of antisocial peer influence and how peer groups may initiate and commit individual crimes. Future research should compare leadership traits of group members during individual criminal incidents and group members who engage in delinquent behavior over multiple incidents.

Relatedly, the present research is relying upon the leadership classification made by court actors (e.g., police, lawyers) and subsequently by two research assistants, introducing two possible sources of bias. First, the leadership classification by court actors is made based upon the behaviors exhibited by the group members, such as selecting the victim, creating the initial plan, and/or disposing of a body (Porter & Alison, 2001, 2006). This leadership classification itself is likely not immune to bias. To the extent that there is bias in the process court actors use to determine leadership, or that court actors were influenced by the facial appearance of the group members in the same way our participants were when making this leadership classification, some of our conclusions would be undermined. Ironically, if court actors were biased by faces in a manner consistent with those of our participants, this would still be consistent with the conclusions of the present research that facial appearance is an extralegal factor in courtroom outcomes. Ultimately, we note that being classified as a leader was strongly associated with very serious sentencing decisions, providing some validity evidence for this approach.

Second, although the coders were blind to targets' actual leadership and group membership, they were not blind to the broad purpose of the studies (i.e., that appearance might be associated with leadership and outcomes). Mitigating this concern is that we had no hypotheses regarding which specific appearance-based traits might separate leaders from followers. Nonetheless, increasing the objectivity of leadership classification in studies and in the courtroom would be preferable to reduce sources of bias.

It is also possible that group members would identify a different group member as a leader or report group leadership to be dynamic. Group members may report leadership during a criminal incident as exclusive to that particular incident. As previously noted, comparisons of group leadership across individual and repeated incidents would elucidate whether leadership-specific traits (i.e., Dominance), hold across multiple contexts. Despite the possibility that group members perceive leadership differently than do court actors, sentencing decisions should, theoretically, be influenced by the group member's role in the crime(s) at hand and not their relative role in the group more broadly.

Finally, groups were identified by newspaper coverage. It is possible that severe, violent, group-perpetrated crimes, such as those represented here, are qualitatively different from minor crimes committed in groups in which leader and follower dynamics may be less apparent. Furthermore, newspapers did not always report on sentencing outcomes or case conclusions. Future research might examine our conclusions among a broader sampling of types of offenses.

CONCLUSION

In summary, the facial appearance of youth in a group-perpetrated criminal incident was possibly associated with leadership status, significantly predicted perceptions of leadership, and influenced subsequent sentencing outcomes. Our findings are consistent with previous literature on facial cues and leadership status. At the same time, to our knowledge the current study is the first to report that facial appearance in sentencing outcomes is an extralegal variable relevant to youth in conflict with the law. Future research should consider how group dynamics and leadership differs depending on the type of crime and if individuals use different facial cues to select leaders when given information about the crime in question.

AUTHOR CONTRIBUTIONS

J.E.S., A.M.C., and E.H. designed the studies. J.E.S. and A.M.C. collected the data. J.E.S. and E.H. performed the analyses. All authors wrote the manuscript.

ORCID iD

Jessica E. Sutherland  <https://orcid.org/0000-0003-2081-8398>

REFERENCES

- Akers, R. L., Krohn, M. D., Lanza-Kaduce, L., & Radosevich, M. (1979). Social learning and deviant behavior: A specific test of a general theory. *American Sociological Review, 44*(4), 635–655.
- Albert, D., Chein, J., & Steinberg, L. (2013). The teenage brain: Peer influences on adolescent decision making. *Current Directions in Psychological Science, 22*(2), 114–120. <https://doi.org/10.1177/0963721412471347>
- Allen, J. P., Chango, J., Szwedo, D., Schad, M., & Marston, E. (2012). Predictors of susceptibility to peer influence regarding substance use in adolescence. *Child Development, 83*(1), 337–350. <https://doi.org/10.1111/j.1467-8624.2011.01682.x>
- Amemiya, J., Monahan, K. C., & Cauffman, E. (2016). Leaders and followers in juvenile offending: Distinguishing correlates and adjustment to incarceration. *Criminal Justice and Behavior, 43*, 899–922. <https://doi.org/10.1177/0093854815622660>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Brown, B. B. (2004). Adolescents' relationships with peers. In R. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 363–394). Wiley.
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. University of California Press.
- Camilletti, C. R., & Scullin, M. H. (2012). Attorney and lay beliefs about factors affecting jurors' perceptions of juvenile offender culpability. *Psychology, Crime & Law, 18*, 113–128. <https://doi.org/10.1080/1068316X.2011.613390>

- Carrington, P. J. (2009). Co-offending and the development of the delinquent career. *Criminology*, 47(4), 1295–1329. <https://doi.org/10.1111/j.1745-9125.2009.00176.x>
- Cherulnik, P. D., Turns, L. C., & Wilderman, S. K. (1990). Physical appearance and leadership: Exploring the role of appearance-based attribution in leader emergence. *Journal of Applied Social Psychology*, 20, 1530–1539.
- Choukas-Bradley, S., Giletta, M., Cohen, G. L., & Prinstein, M. J. (2015). Peer influence, peer status, and prosocial behavior: An experimental investigation of peer socialization of adolescents' intentions to volunteer. *Journal of Youth and Adolescence*, 44(12), 2197–2210. <https://doi.org/10.1007/s10964-015-0373-2>
- da Silva, T., Woodhams, J., & Harkins, L. (2018). "An adventure that went wrong": Reasons given by convicted perpetrators of multiple perpetrator sexual offending for their involvement in the offense. *Archives of Sexual Behavior*, 47(2), 443–456. <https://doi.org/10.1007/s10508-017-1011-8>
- Dmitrieva, J., Gibson, L., Steinberg, L., Piquero, A., & Fagan, J. (2014). Predictors and consequences of gang membership: Comparing gang members, gang leaders, and non-gang-affiliated adjudicated youth. *Journal of Research on Adolescence*, 24(2), 220–234. <https://doi.org/10.1111/jora.12111>
- Dryer, D. C. & Horowitz, L. M. (1997). When do opposites attract? Interpersonal complementarity versus similarity. *Journal of Personality and Social Psychology*, 72(3), 592–603.
- Faucher, C. (2009). Fear and loathing in the news: A qualitative analysis of Canadian print news coverage of youthful offending in the twentieth century. *Journal of Youth Studies*, 12, 439–456. <https://doi.org/10.1080/1367620902897426>
- Flowe, H. D. (2012). Do characteristics of faces that convey trustworthiness and dominance underlie perceptions of criminality? *PLOS ONE*, 7, Article e37253. <https://doi.org/10.1371/journal.pone.0037253>
- Forbes, E. E., & Dahl, R. E. (2010). Pubertal development and behavior: Hormonal activation of social and motivational tendencies. *Brain and Cognition*, 72(1), 66–72. <https://doi.org/10.1016/j.bandc.2009.10.007>
- Geen, R. G. (1991). Social motivation. *Annual Review of Psychology*, 42, 377–389. <https://doi.org/10.1146/annurev.ps.42.020191.002113>
- Grabo, A., & Van Vugt, M. (2018). Voting for a male warrior or female peacekeeper? Testing the evolutionary contingency hypothesis in the 2016 U.S. presidential elections. *Evolutionary Psychology*, 16(2), 1–9. <https://doi.org/10.1177/1474704918773267>
- Graham, J. R., Harvey, C. R., & Puri, M. (2017). A corporate beauty contest. *Management Science*, 63(9), 3044–3056.
- Haynie, D. L., Doogan, N. J., & Soller, B. (2014). Gender, friendship networks, and delinquency: A dynamic network approach. *Criminology*, 52(4), 688–722. <https://doi.org/10.1111/1745-9125.12052>
- Haynie, D. L., & Osgood, D. W. (2005). Reconsidering peers and delinquency: How do peers matter? *Social Forces*, 84(2), 1109–1130. <https://doi.org/10.1353/sof.2006.0018>
- Hehman, E., Flake, J. K., & Freeman, J. B. (2018). The faces of group members share physical resemblance. *Personality and Social Psychology Bulletin*, 44(1), 3–15. <https://doi.org/10.1177/0146167217722556>
- Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2015). Picking teams: When dominant facial structure is preferred. *Journal of Experimental Social Psychology*, 59, 51–59.
- Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, 49, 747–754.
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103(1), 54–69.
- Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, 98(2), 155–163. <https://doi.org/10.1037/0033-295X.98.2.155>
- LaFontana, K. M., & Cillessen, A. H. N. (2010). Developmental changes in the priority of perceived status in childhood and adolescence. *Social Development*, 19(1), 130–147. <https://doi.org/10.1111/j.1467-9507.2008.00522.x>
- Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior*, 28, 18–27.
- Malloy, L., Cauffman, E., Kimonis, E., & Kelly, E. (2007, March). *Peers and adjustment to incarceration* [Paper presentation]. Annual Meeting of the Society for Research on Child Development, Boston, MA.
- McCoach, D. B., & Adelson, J. L. (2010). Dealing with dependence (Part I): Understanding the effects of clustered data. *Gifted Child Quarterly*, 54(2), 152–155.
- McGloin, J. M., & Nguyen, H. (2012). It was my idea: Considering the instigation of co-offending. *Criminology*, 50, 463–494. <https://doi.org/10.1111/j.1745-9125.2011.00266.x>
- McGloin, J. M., & Rowan, Z. R. (2015). A threshold model of collective crime. *Criminology*, 53(3), 484–512. <https://doi.org/10.1111/1745-9125.12077>
- McGloin, J. M., & Thomas, K. J. (2016). Incentives for collective deviance: Group size and changes in perceived risk, cost, and reward. *Criminology*, 54(3), 459–486. <https://doi.org/10.1111/1745-9125.12111>
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100, 674–701.

- Monahan, K. C., Steinberg, L., & Cauffman, E. (2009). Affiliation with antisocial peers, susceptibility to peer influence, and antisocial behavior during the transition to adulthood. *Developmental Psychology, 45*(6), 1520–1530. <https://doi.org/10.1037/a0017417>
- Olivola, C. Y., Eubanks, D. L., & Lovelace, J. B. (2014). The many (distinctive) faces of leadership: Inferring leadership domain from facial appearance. *The Leadership Quarterly, 25*, 817–834. <https://doi.org/10.1016/j.leaqua.2014.06.002>
- Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences, 18*(11), 566–570. <https://doi.org/10.1016/j.tics.2014.09.007>
- Olivola, C. Y., & Todorov, A. (2010a). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior, 34*, 83–110. <https://doi.org/10.1007/s10919-009-0082-1>
- Olivola, C. Y., & Todorov, A. (2010b). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology, 46*, 315–324.
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America, 105*(32), 11087–11092. <https://doi.org/10.1073/pnas.0805664105>
- Porter, L. E., & Alison, L. J. (2001). A partially ordered scale of influence in violent group behavior: An example from gang rape. *Small Group Research, 32*(4), 475–497. <https://doi.org/10.1177/104649640103200405>
- Porter, L. E., & Alison, L. J. (2006). Leadership and hierarchies in criminal groups: Scaling degrees of leader behaviour in group robbery. *Legal and Criminological Psychology, 11*(2), 245–265. <https://doi.org/10.1348/135532505X68692>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). SAGE.
- Re, D. E., Hunter, D. W., Coetzee, V., Tiddeman, B. P., Xiao, D., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2013). Looking like a leader—facial shape predicts perceived height and leadership ability. *PLOS ONE, 8*, Article e80957. <https://doi.org/10.1371/journal.pone.0080957>
- Re, D. E., & Rule, N. O. (2017). Distinctive facial cues predict leadership rank and selection. *Personality and Social Psychology Bulletin, 43*, 1311–1322. <https://doi.org/10.1177/0146167217712989>
- Reiss, A. J. (1988). Co-offending and criminal careers. *Crime and Justice, 10*, 117–170. <https://doi.org/10.1086/449145>
- Rockloff, M. J., & Dyer, V. (2007). An experiment on the social facilitation of gambling behavior. *Journal of Gambling Studies, 23*(1), 1–12. <https://doi.org/10.1007/s10899-006-9042-4>
- Rovner, J. (2017, October 13). *Juvenile life without parole: An overview*. The Sentencing Project. <http://www.sentencingproject.org/publications/juvenile-life-without-parole/>
- Rowan, Z. R., McGloin, J. M., & Nguyen, H. (2018). Capitalizing on criminal accomplices: Considering the relationship between co-offending and illegal earnings. *Justice Quarterly, 35*(2), 280–308. <https://doi.org/10.1080/07418825.2017.1308540>
- Ruigrok, N., van Atteveldt, W., Gagestein, S., & Jacobi, C. (2017). Media and juvenile delinquency: A study into the relationship between journalists, politics, and public. *Journalism, 18*, 907–925. <https://doi.org/10.1177/1464884916636143>
- Shi, B., & Xie, H. (2012). Socialization of physical and social aggression in early adolescents' peer groups: High-status peers, individual status, and gender. *Social Development, 21*(1), 170–194. <https://doi.org/10.1111/j.1467-9507.2011.00621.x>
- Stewart, J. E. (1980). Defendant's attractiveness as a factor in the outcome of criminal trials: An observational study. *Journal of Applied Social Psychology, 10*, 348–361.
- Stolzenberg, L., & D'Alessio, S. J. (2008). Co-offending and the age-crime curve. *Journal of Research in Crime & Delinquency, 45*, 65–86. <https://doi.org/10.1177/0022427807309441>
- Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition, 127*(1), 105–118. <https://doi.org/10.1016/j.cognition.2012.12.001>
- Thornberry, T. P. (1987). Toward an interactional theory of delinquency. *Criminology, 25*(4), 863–892. <https://doi.org/10.1111/j.1745-9125.1987.tb00823.x>
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*, 1623–1626. <https://doi.org/10.1126/science.1110589>
- 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*(1), 519–545. <https://doi.org/10.1146/annurev-psych-113011-143831>
- van Mastrigt, S. B., & Farrington, D. P. (2011). Prevalence and characteristics of co-offending recruiters. *Justice Quarterly, 28*(2), 325–359. <https://doi.org/10.1080/07418825.2010.482535>
- Van Vugt, M., & Grabo, A. E. (2015). The many faces of leadership: An evolutionary-psychology approach. *Current Directions in Psychological Science, 24*, 484–489. <https://doi.org/10.1177/0963721415601971>
- Warr, M. (1996). Organization and instigation in delinquent groups. *Criminology, 34*, 11–37. <https://doi.org/10.1111/j.1745-9125.1996.tb01193.x>
- Weerman, F. M., Lovegrove, P. J., & Thornberry, T. (2015). Gang membership transitions and its consequences: Exploring changes related to joining and leaving gangs in two countries. *European Journal of Criminology, 12*(1), 70–91. <https://doi.org/10.1177/1477370814539070>

- Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science, 26*, 1325–1331. <https://doi.org/10.1177/0956797615590992>
- Wong, I. W. J., Toh, D. P., Hung, P. P. L., & Ang, R. P. (2013). Delinquency in gangs—selection or socialization? *Aggression and Violent Behavior, 18*(6), 784–791. <https://doi.org/10.1016/j.avb.2013.10.001>
- Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior, 15*, 603–623. <https://doi.org/10.1007/BF01065855>
- Zebrowitz, L. A., & Montepare, J. M. (2005). Appearance DOES matter. *Science, 308*, 1565–1566.

Jessica E. Sutherland, PhD, is an assistant professor in the Department of Psychology at York University in Toronto, Canada. Her research has contributed to the fields of juvenile offending, developmental science and the law, and adolescent wellbeing.

Ana M. Cojocariu, BA, is a graduate student in the Department of Psychology at Ryerson University in Toronto, Canada. Her research explores intergroup contact, political psychology, and prejudice directed at marginalized groups.

David M. Day, PhD, C. Psych, is a professor in the Department of Psychology at Ryerson University in Toronto, Canada. His research explores delinquency and offending, most recently examining the developmental trajectories of youth in conflict in the law.

Eric Hehman, PhD, is an assistant professor in the Department of Psychology at McGill University in Montreal, Canada. His research examines how individuals perceive and evaluate one another across group boundaries, and the downstream consequences of impression formation processes.