Running head: PROBLEM BEHAVIOR AND FACIAL TRUSTWORTHINESS

Prospective Associations Between Boys' Substance Use and Problem Behavior Histories and

Their Facial Trustworthiness in Adulthood

Zoe M. Alley

School of Psychological Science

Oregon State University

David C. R. Kerr

School of Psychological Science

Oregon State University

and Oregon Social Learning Center

John Paul Wilson

Department of Psychology

Montclair State University

Nicholas O. Rule

Department of Psychology

University of Toronto

Correspondence concerning this article should be addressed to Zoe M. Alley, School of

Psychological Science, Oregon State University, 2950 SW Jefferson Way, Corvallis, OR 97331,

USA. Email: alleyz@oregonstate.edu Telephone: 541.737.1364

Abstract

Introduction: People whose faces look untrustworthy tend to receive harsher social evaluations, including more severe criminal sentences. Yet little is known about how much facial trustworthiness reflects individuals' behavioral histories. We examined whether adolescent histories of delinquency and substance use predict strangers' perceptions of young men's facial trustworthiness. **Methods:** Boys ($n = 206$) recruited from schools with higher juvenile crime rates were assessed repeatedly from ages 10-24 years, including arrest records and self-reported delinquency and substance use. Coders blind to the study's purpose rated participants' facial trustworthiness from photographs taken at ages 14 and 24; parent-reported childhood family income and coder ratings of attractiveness and positive affect at age 24 were considered as controls. **Results:** Facial trustworthiness at age 24 (but not age 14) negatively correlated with all measures of problem behavior. Yet, self-reported tobacco use occasions from ages 12-23 had the strongest association with facial trustworthiness at age 24, a relation that persisted when controlling for arrests and delinquency from ages 12-23, other substance use, family income, ratings of age-24 positive facial affect, attractiveness, and age-14 facial trustworthiness ($\beta = -.29$, 95% CI [-.42, -.15], $p < .001$). **Discussion:** Although boys' early facial trustworthiness did not relate to their later problem behavior, men with histories of more delinquency and tobacco use appeared less facially trustworthy as adults. Appearance-related biases may have forensic and healthcare implications for young men. Additionally, prevention efforts could leverage information about the early impacts of tobacco use on appearance.

Key words: adolescence, Dorian Gray effect, face perception, tobacco, trustworthiness

Prospective Associations Between Boys' Substance Use and Problem Behavior Histories and

Their Facial Trustworthiness in Adulthood

People readily infer others' personality characteristics from nonverbal information, such

as body movement and eye contact (Ambady & Skowronski, 2008). Yet a person's face is often

the most immediate source of nonverbal cues. Consequently, the face can direct interpersonal

interactions via overt structural and dynamic physiognomic cues (e.g., emotional expressions).

Indeed, people spontaneously infer personality traits from more stable facial characteristics, such

as attractiveness or trustworthiness, with as little as a 50-ms glance (Re & Rule, 2015; Todorov,

Pakrashi, & Oosterhof, 2009).

The inferences people quickly make about others' facial trustworthiness can have

significant social consequences (Wilson & Rule, 2017). For example, people tend to invest more

money with individuals who have trustworthy-looking faces (Van't Wout & Sanfey, 2008) and

trustworthy-looking people are more likely to receive loans from others (Duarte et al., 2012).

Expert nurses show similar evidence of such "face-ism," reporting more inclination to care for

trustworthy- than untrustworthy-looking patients (Mattarozzi et al., 2017). In addition, inferences

of trustworthiness based on physical appearance can unfairly bias criminal sentencing decisions.

Specifically, among men convicted of murder, those who appeared less trustworthy to lay raters

were more likely to have received the death sentence—even in cases in which the men had been

exonerated (Wilson & Rule, 2015, 2016).

Although research indicates that facial trustworthiness has the potential to influence a

broad array of social interactions and life outcomes, most studies focus on the behaviors and

decisions of the people perceiving another person's face. In the present study, we focus instead

on the extent to which individuals' facial trustworthiness relates to their own past and future behavior.

What is a trustworthy face?

Several physical characteristics are associated with perceptions of trustworthiness. Faces people find trustworthy tend to have upturned lips and raised brows at rest—features that resemble a smiling or approachable facial expression. Conversely, low brows and downturned lips (similar to the emotional expression of anger) characterize the typical "untrustworthy" face (Ma et al., 2015; Oosterhof & Todorov, 2008; Todorov, Baron, & Oosterhof, 2008). Faces may thus appear trustworthy or untrustworthy because of their similarity to relevant emotions—an emotion-overgeneralization effect (Zebrowitz & Montepare, 2008). Although trustworthiness judgments may therefore be sensitive to transient emotional expressions (i.e., people look more trustworthy when smiling), variation in facial trustworthiness perceptions are still detectable among smiling individuals and rely on more than facial affect alone (Nurmoja & Bachman, 2014). Trustworthiness also correlates with physical attractiveness, though the two remain meaningfully distinct (e.g., Sofer, Dotsch, Wigboldus, & Todorov, 2015). Thus, although these other aspects of the face can influence trustworthiness perceptions, facial trustworthiness varies separately and forms a distinct dimension of face perception.

Lay notions of facial appearance typically characterize evaluations of trustworthiness and attractiveness as highly subjective, largely residing "in the eye of the beholder." However, many studies have demonstrated that ratings of facial trustworthiness quickly converge across perceivers, allowing researchers to calculate mean trustworthiness ratings with acceptable internal consistency from relatively small numbers of raters (Todorov et al., 2009; Willis & Todorov, 2006). Whereas substantial variability exists in different individuals' judgments of

faces, people generally agree about who looks attractive or trustworthy. Remarkably, these judgments also do not seem to vary much according to the raters' demographic characteristics. For example, children and adults from Western and Eastern countries utilize the same facial characteristics to judge facial trustworthiness (Cogsdill, Todorov, Spelke, & Banaji, 2014; Ma, Xu, & Luo, 2016). Despite some evidence that women may rate female faces as more trustworthy than men do, the same gender difference has not emerged for male faces (Mattarozzi, Todorov, Marzocchi, Vicari, & Russo, 2015). The initial impressions that convenience samples of lay raters form of individuals' faces may therefore represent the initial impressions that others in their lives also make. Here, we focus on whether such impressions might reflect or influence those individuals' behaviors.

Does facial trustworthiness match trustworthy behavior?

Trustworthiness judgments constitute a basic dimension of person perception: assessing a target's approachability versus threat (Oosterhof & Todorov, 2008). Quick and automatic, these judgments have consequences in lab tasks and relate to meaningful real world outcomes (e.g., Wilson & Rule, 2016). These findings raise the important question of whether facial trustworthiness is a valid indicator of trustworthy behavior.

Interest in the validity of inferences of facial trustworthiness has recently grown (Wilson & Rule, 2017). Backed by empirical demonstrations that facial trustworthiness perceptions do not predict behavior (e.g., cheating on a test), some researchers have concluded that trustworthiness inferences are not valid (e.g., Rule, Krendl, Ivcevic, & Ambady, 2013; Todorov, Olivola, Dotsch, & Mende-Siedlicki, 2015). But other researchers have determined that trustworthiness inferences may contain a "kernel of truth" (Bonnefon, Hopfensitz, & De Neys, 2015). Although few studies have investigated the validity of impressions based on facial

trustworthiness, evidence suggests that judgments regarding related facial features are accurate in some cases. For example, men with higher facial width to height ratios were less likely to be trusted with money in an economic game by peers who viewed only their faces and were, in fact, less likely to behave cooperatively than other men (Stirrat & Perrett, 2010). In another study, raters could distinguish between photographs of violent and nonviolent criminals, indicating that valid facial cues to violence (an inherently untrustworthy behavior) may exist (Stillman, Maner, & Baumeister, 2010). Other work found that people correctly trusted reciprocators more often than abusers in a trust game but only when judgments were effortless and relied less on conscious trustworthiness judgments (Bonnefon, Hopfesnitz, & De Neys, 2013). Whether facial trustworthiness itself relates to the likelihood of engaging in stereotypically untrustworthy future behaviors requires study, however. Several plausible explanations may support an association between one's facial trustworthiness and (un)trustworthy behavior.

First, past work suggests that untrustworthy behaviors could lead to an untrustworthy face. Behaviors consistent with psychological traits may spur changes in physical appearance that match stereotypes associated with those characteristics—a Dorian Gray effect, so named for the titular protagonist of Oscar Wilde's novel, whose hideous deeds manifested in a grotesque appearance (Zebrowitz, 1997; Zebrowitz et al., 1998). For instance, Malatesta, Fiore, and Messina (1987) found that faces can become congruent with personality over time. In the case of antisocial traits, a person with violent tendencies and behaviors may make facial expressions indicative of anger more often than someone who does not have such tendencies. These facial expressions may shape the creasing and musculature of the face such that the person appears increasingly negative and threatening even at rest (e.g., Adams, Garrido, Albohn, Hess, & Kleck, 2016; Bjornsdottir & Rule, 2017).

Second, an untrustworthy face could lead to more untrustworthy behaviors through expectancy effects, whereby individuals' inferences about another person's personality affects how that person behaves towards him or her (see Gilbert, 1995). Facial cues may therefore elicit reactions from the social environment that facilitate confirmation biases from social partners (Zebrowitz, 1997). More recent work found that face-based judgments of participants predicted trustworthy behavior, and that the targets' expectations about how others might see them mediated the link between the judgments and their behaviors (Slepian & Ames, 2015). In this way, an association between facial characteristics and behavior may reflect changes that occurred in behavior across time, even if the person's apparent facial trustworthiness has remained stable.

"Common causes" are a third explanation for why facial trustworthiness and trustworthy behavior could be associated; that is, individual and contextual factors might explain both facial characteristics and the behaviors congruent with perceptions of those characteristics (Zebrowitz, 1997). For example, socioeconomic disadvantage is associated with adolescents' problem behaviors, which are considered untrustworthy (Ary et al., 1999). This same disadvantage could relate to the development of an untrustworthy appearance. But because such biases can reflect spurious or imperfect trait-behavior associations, some individuals may find themselves with an undeservedly untrustworthy appearance.

The Present Study

Our primary question here concerned whether a community sample of at-risk boys who engaged in more untrustworthy problem behaviors during adolescence (i.e., delinquency) and early adulthood would be perceived by strangers—who knew nothing of their histories—as less trustworthy at age 24. Furthermore, by accounting for facial trustworthiness during early

adolescence, we could test the Dorian Gray hypothesis that an association between earlier delinquency and later facial trustworthiness reflects changes in the face across time. This analysis also permits testing the other two paths described above (expectancy and common causes) by evaluating whether early facial trustworthiness or another early contextual factor—family income—predicts later delinquency or partially explains the association between delinquency and later facial trustworthiness. To our advantage, these behaviors were all measured prospectively and reflect the boys' and men's "real life" behaviors outside of a laboratory context.

Facial trustworthiness and substance use. Unlike delinquency, substance use is not inherently "untrustworthy" behavior, But delinquency and substance use are closely related in adolescence (Jessor, 1991). Thus, it seems likely that if adolescent delinquency relates to facial trustworthiness, then substance use will too. This latter association may simply reflect a confound; or, it may mechanistically link delinquency and facial appearance, consistent with the Dorian Gray effect.

Most adolescents do not use alcohol, marijuana, or other drugs intensively enough to plausibly expect the substances to affect their physical appearance by early adulthood. But tobacco use could have such effects. Indeed, smoking is known to affect physical appearance, including the skin (e.g., premature aging, acne, wound healing; Freiman et al., 2004; Morita, 2007; Schafer et al., 2001), teeth, general appearance of health (e.g., skin tone; effects of weight/appetite suppression), and individual facial features (e.g., Okada et al., 2013), including those implicated in facial trustworthiness judgments (e.g., the lips; Ma et al., 2015; Oosterhof & Todorov, 2008; Todorov et al., 2008). Thus, we had some reason to expect men's tobacco use histories to be associated with strangers' perceptions of their faces.

Given that antisocial behavior and socioeconomic disadvantage both relate to smoking and to other poor health behaviors that influence facial characteristics (Cohen et al., 2010; Odgers et al., 2008), and because delinquency and deviant peer affiliation occur more commonly among tobacco-using adolescents than others (Ary et al., 1999; Watts & Wright, 1990), tobacco use may explain an association between problem behaviors and changes in facial trustworthiness. Thus, we examined the association that boys' substance use had with their facial trustworthiness in adulthood and explored whether their tobacco use, in particular, relates to trustworthiness unique from delinquency, other substance use, and socioeconomic status.

Hypotheses. Although we know of no longitudinal research regarding the validity of facial trustworthiness or its development across adolescence and early adulthood, the literature previewed thus far supports the following hypotheses. First, consistent with the Dorian Gray effect, we predicted that boys' delinquency in adolescence and early adulthood would relate to lower facial trustworthiness at age 24; that is, that facial trustworthiness may follow from the participants' behavioral histories. Consistent with the expectancy effects and shared causes mechanisms, we also explored whether their delinquency would relate to early facial trustworthiness; that is, whether facial trustworthiness during early adolescence relates to participants' future behavior. Next, we predicted that substance use (tobacco use, in particular) would uniquely relate to the association between men's histories of delinquency and their adult facial trustworthiness. By accounting for early facial trustworthiness in these models, we planned to evaluate the extent to which facial trustworthiness predated (vs. followed) the boys' behavioral histories. All models accounted for the potential confounding roles of childhood family income and adolescent substance use (Ary et al., 1999; Watts & Wright, 1990) and

statistically adjusted for the expected positive associations between trustworthiness and positive

affect. Given its association with facial trustworthiness, we also adjusted for facial attractiveness.

Method

Participants

Participants were drawn from the Oregon Youth Study (OYS), a longitudinal study of

206 boys recruited in 1983–1985 designed to examine the development of antisocial behavior

(see Capaldi & Patterson, 1989). After obtaining ethics approval from the Oregon Social

Learning Center IRB, the original investigators selected the schools with the highest rates of

police-reported delinquent episodes by juveniles within a medium sized metropolitan area. OYS

staff attempted to recruit boys in entire fourth-grade classes in those schools. Parents of 74% of

the targeted boys agreed to allow their son to participate (Capaldi & Patterson, 1989). Study staff

received parental consent for participants when they were minors and participants' own consent

thereafter; staff renewed informed consent with participants regularly across the longitudinal

study as new procedures and measures were used. Compensation was key to full participation

and retention (Capaldi, Chamberlain, Fetrow, & Wilson, 1997); participants and their parents

each were provided $100 for participating in annual interviews (i.e., up to $300 per family), plus

additional compensation for tasks such as home observations. The sample was primarily White[1]

---

[1] White participants ($n = 157$) had higher parental incomes, $t(30.79,$ unequal variances$) = -3.25$,

$p = .003$, $r_{\text{effect size}} = -.36$, and (marginally) lower cumulative arrest scores than participants of

another race/ethnicity ($n = 20$), $t(175) = 1.99$, $p = .05$, $r_{\text{effect size}} = .23$, but did not differ on any

other predictor, $|t(175)|s = 0.52–1.77$, $ps = .60–.08$, $|r_{\text{Seffect size}}| = .06–.21$, including age-24

trustworthiness, $t(175) = -0.21$, $p = .84$, $r_{\text{effect size}} = -.02$; age-24 attractiveness, $t(175) = -0.70$, $p =$

(90%; 3% African American, 2% American Indian, 1% Mexican American, and 5% other

identities) and generally consisted of individuals from low socioeconomic status families

(median annual income at study entry = $15,000). Multimethod, multiagent assessment of the

participants occurred almost annually from ages 10–24 years. Facial photographs taken at age 24

were available for 177 of the 206 participants. Facial photographs taken at age 14 were available

for 183 of the participants, but for only 159 of the 177 individuals who also had age-24 photos.

Note that maximum likelihood estimation allowed for the inclusion of all 177 individuals who

had age-24 photographs in the sample even if age-14 photographs were missing, as stated below.

Behavioral Measures

  Childhood family income. Parents reported their annual income when the participants

were ages 10, 11, and 12 years; the mean of these *Z*-standardized variables served as a control

variable.

  Delinquency/Arrest History. Participants completed the Elliot Delinquency Scale

(Elliott et al., 1983) annually from age 13 (when it was first administered) to 23 years. This self-

report measure was designed as an analog to the Federal Bureau of Investigation's Uniform

Crime Reports arrest measure. Participants reported frequencies with which they engaged in a

range of antisocial behaviors during the prior year (e.g., theft, vandalism, and violence). OYS

investigators capped responses to each item at 365 to reduce skew and the influence of specific

years on the total score. Previous research has established the internal consistency (across waves,

---

.48, $r_{\text{effect size}}$ = -.08; age-14 trustworthiness, $t(157)$ = -0.25, p = .80, $r_{\text{effect size}}$ = -.02; or tobacco-

use occasions, $t(175)$ = -0.41, p = .68, $r_{\text{effect size}}$ = -.05. Thus, we did not adjust for race/ethnicity

in the regressions.

PROBLEM BEHAVIOR AND FACIAL TRUSTWORTHINESS

mean $\alpha = .76$; Kerr et al., 2011), reliability, and validity of this measure (Elliott et al., 1983). We applied a logarithmic transformation to reduce skew, kurtosis, and the influence that unusually high scores at a single wave could have on the participants' lifetime means before calculating the mean of the transformed scores across waves for each participant.

OYS staff collected annual arrest counts from public records for each participant. Charges related to traffic violations (e.g., driving uninsured, speeding tickets), protective custody, offenses against the participants, contempt of court, or uncodeable offenses were excluded (Wiesner et al., 2007). We log-transformed the total number of lifetime charges through age 23 years for the same reasons listed above.

The self-reported delinquency and arrest variables correlated moderately, $r(176) = .48$, $p < .001$. We therefore averaged each participant's standardized scores to create a delinquency/arrests variable.

Tobacco-use occasions. Starting at age 12, participants were annually asked to self-report the number of times (to a maximum of 999) they used tobacco in the previous year. We summed these reports across ages 12–23 years, dividing the sums by 1,000 so that the parameter estimates approximated the scale of the other variables; thus, the tobacco-use variable reported in Table 1 approximates the cumulative number of tobacco-use occasions in thousands.

Non-tobacco substance use history. Self-reported frequency of tobacco, marijuana, and alcohol use were assessed annually from ages 12–23 years on a 9-point scale from 0 to 8. These values signified the frequency of substance use in the previous year (0 = *never*, 1 = *once or twice*, 2 = *every 2-3 months*, 3 = *once a month*, 4 = *every 2-3 weeks*, 5 = *once a week*, 6 = *2-3 times a week*, 7 = *once a day*, 8 = *2-3 times a day*). Annual frequencies of "other drug" use were collected from ages 12–23, except at ages 19 and 20. Mean frequencies across assessment years

were calculated for marijuana, alcohol, and other drug use. To minimize the number of model predictors (and because the means significantly correlated: rs = .37-.69, all ps < .01; Table 1), we calculated a nontobacco substance-use control variable by standardizing the scores for each variable and calculating the mean of the Z-scores (α = .74).

Measures of Facial Characteristics

Participants were photographed at multiple assessment waves. All ratings of participants' faces were made by ethics-trained research assistants (RAs), as the Oregon Social Learning Center's Internal Review Board required to protect the participants' confidentiality. RAs were blind to the participants' behavioral histories. We emphasize that it was the participants (i.e., OYS males) and not the RA raters who constituted the unit of analysis in this study and thus provide the main source of statistical power. We therefore sampled enough raters to provide reliable judgments of the targets and used the mean ratings as variables in the analysis (e.g., Schönbrodt & Perugini, 2013; Webb, Shavelson, & Haertel, 2006). As in previous research demonstrating the automatic nature of facial trustworthiness judgments and agreement among untrained raters (Todorov et al., 2009), the facial trustworthiness ratings converged to an acceptable reliability criterion (αs $\geq$ .80.), indicating that mean facial trustworthiness ratings would not change substantially by adding more RA raters to the sample.

Trustworthiness. Nineteen RAs (15 women, 4 men; α = .85) viewed the participants' photographs taken at age 24 in random order on a computer screen and responded to the question, "How trustworthy does this person look?" using a scale from 1 (*Very Untrustworthy*) to 7 (*Very Trustworthy*). RAs were blind to the purpose of the study. Ratings of the age-14 photographs were collected in the same manner from another 19 RAs (13 women, 6 men; α = .83).

PROBLEM BEHAVIOR AND FACIAL TRUSTWORTHINESS

Attractiveness. A subsample of the hypothesis-blind RAs ($N = 16$; 12 women, 4 men) also rated men's attractiveness in the age-24 photographs from 1 (*Very Unattractive*) to 7 (*Very Attractive*) in response to the question, "How attractive does this person look?" As with the facial trustworthiness ratings, these judgments showed satisfactory convergence ($\alpha = .83$), so we used each participant's mean score.

Positive affect. Although participants were instructed to hold neutral expressions when photographed, some smiled. To control for the influence of facial expression on perceptions of trustworthiness at age 24 (Nurmoja & Bachmann, 2014), two raters coded each photo from 1 (*No Positive Affect*) to 5 (*Strong Positive Affect*) with adequate interrater reliability ($\kappa = .96$), similar to past work (Zebrowitz & Montepare, 1992; Zebrowitz et al., 1993). Their mean scores served as a control variable in the analysis.

Missing Data

Two-thirds or more of the total observations in cross-time variables were present for all but one participant, who was missing four out of ten reports of other drug use. He denied other drug use at all six assessments for which his report was present, so we judged it appropriate to retain his data in the analysis.

Of the original 206 men, 29 did not participate in the age-24 photo collection. However, 25 of these men participated in a subsequent wave; thus, true study attrition was very low. Men who participated at age 24 had histories of less frequent marijuana use ($M = 1.05$, $SD = 1.28$) than those who did not participate at age 24 ($M = 1.65$, $SD = 1.50$), $t(202) = -2.19$, $p = .029$, $r_{\text{effect size}} = -.15$, but did not significantly differ on any other behavioral predictor variables, all $|t|\text{s} \leq 1.52$, $p\text{s} \geq .13$, $|r_{\text{Seffect size}}| < .16$. Photographs at both ages 13 and 24 were available for 159 participants, as noted above.

PROBLEM BEHAVIOR AND FACIAL TRUSTWORTHINESS

Statistical analyses

We tested hypotheses via linear regression using *MPlus* (Muthén & Muthén, 1998-2011). As covariances were specified among predictor variables, our use of maximum likelihood estimation included data from all 177 men with age-24 photographs in the model even if their age-14 photographs were unavailable.

Results

Descriptive Statistics

Table 1 depicts the variability in delinquency, arrests, substance use, and facial characteristics of the men in the sample. The average number of arrests was approximately four; about 64% of the men were arrested at least once before age 24 years. Means of total tobacco-use occasions reflected a range of exposures ($M = 3,430$, $SD = 3,540$; *Range:* 0-10,990). Of the 177 participants, 109 smoked at least once per day at some time point before age 24. Averaged across all waves, participants reported using alcohol about every 2-3 months, smoking marijuana about once or twice in the past year, and rarely using other drugs. Note that these averages included early time points when substance use is less common (e.g., ages 12, 13). The average facial trustworthiness rating was in the approximate middle of the scale at both ages 13 and 24, and the average rating for age-24 attractiveness was near the lower end of the scale. Positive affect ranged widely, from no smiles to full smiles.

Correlations

Facial trustworthiness at age 24 was significantly associated with prior delinquency and arrest history, and with facial trustworthiness ratings at age 14. However, age 14 facial trustworthiness did not significantly correlate with delinquency or arrests.

Similarly, facial trustworthiness at age 24 (but not at age 14) correlated with every substance use variable (except alcohol use); age 24 attractiveness correlated with tobacco use; and family income negatively correlated with delinquency, marijuana, and tobacco use. All measures of substance use and delinquency were significantly intercorrelated (e.g., other drug use and delinquency). Thus, we controlled for the substance use variables as potential confounds via multiple regression to investigate the potential association of delinquent behavior with later facial trustworthiness.

Regression Predicting Age-24 Facial Trustworthiness

We regressed facial trustworthiness at age 24 years on delinquency/arrests, cumulative tobacco use, nontobacco substance use, childhood family income, ratings of positive facial affect, and early adolescent facial trustworthiness (Table 2). To allow for maximum likelihood estimations of missing data, all predictors measured before age 24 (i.e., all but positive affect) were allowed to covary (see table note).

Surprisingly, delinquency/arrests did not independently predict later facial trustworthiness. Tobacco use occasions, on the other hand, significantly related to lower ratings of facial trustworthiness at age 24, $\beta = -.29$, 95% CI [-.42, -.15), $p < .001$. Of note, parent income also predicted age 24 facial trustworthiness even when controlling for age 14 facial trustworthiness. Thus, greater tobacco use and lower parent income predicted decreases in facial trustworthiness by adulthood.

We then tested whether an association between tobacco use and attractiveness better explained the relations between tobacco use and facial trustworthiness, given that facial trustworthiness and attractiveness correlate. After adding age-24 attractiveness as a control

variable in Model 1 (β = .44, 95% CI [.34, .54], p < .001), tobacco-use occasions (β = -.20, 95% CI [-.32, -.08], p = .001) continued to significantly predict lower age 24 facial trustworthiness.

Discussion

In this prospective study of an at-risk community sample of men, histories of arrest and self-reported delinquent behaviors significantly negatively correlated with strangers' perceptions of facial trustworthiness at adulthood, but not in early adolescence. That is, in apparent consistency with the Dorian Gray effect, decrements in facial trustworthiness followed untrustworthy behavior. The effect of tobacco use overshadowed the association between delinquent behavior and later facial trustworthiness in a regression analysis, however. Contrary to the Dorian Gray hypothesis that boys' misdeeds would manifest in their adult faces, models were more consistent with the interpretation that their tobacco use promoted the changes observed in their facial trustworthiness. Although tobacco use may be a misdeed itself, particularly if it occurs illegally during adolescence, it is likely a correlate rather than core component of conventionally untrustworthy acts.

Yet, in both the present and past studies, tobacco use covaried with delinquency (Ary et al., 1999). Higher levels of tobacco use during adolescence and early adulthood then related to how trustworthy the men looked at age 24 years here. This association held when accounting for a number of individual and contextual factors associated with tobacco use—namely, socioeconomic status and other substance use—that could transmit signals of untrustworthiness to others. Early adolescent facial trustworthiness did not significantly relate to tobacco use, other substance use, or delinquent behavior. Thus, a tendency for untrustworthy-looking boys to use tobacco on more occasions across their early lives does not explain the primary findings. Rather,

tobacco use may contribute over time to the development of a facial appearance deemed untrustworthy by others.

These associations notwithstanding, mechanisms underlying the link between tobacco use and facial trustworthiness require further sensitive investigation. Potential mechanisms could include physical effects of smoke exposure on the teeth, skin (e.g., aging, acne, wound healing; Freiman et al., 2004; Schafer et al., 2001), and overall appearance of health (e.g., given weight and appetite suppression). People who use tobacco might also engage in other adverse health behaviors that affect their facial appearance. Although we included attractiveness and positive affect as control variables, their failure to account for the links between tobacco use and facial trustworthiness potentially eliminates them as mechanisms; for example, it appears unlikely that tobacco use reduces apparent trustworthiness by degrading attractiveness.

The fact that men who engaged in more problem behaviors in adolescence and early adulthood looked less trustworthy at age 24 may suggest a kernel of truth in judgments of their appearance; that is, that untrustworthy-looking people have done untrustworthy things. Notably, however, both tobacco use and facial trustworthiness ratings were far from perfect correlates of antisocial behavior, suggesting that strangers unfairly judged many of the men. Thus, tobacco use may influence some people's appearance in a way that encourages false impressions of delinquent traits. Such misperceptions may have important life-course implications for tobacco users, as facial trustworthiness can meaningfully influence important outcomes, such as decisions about suspected criminals' honesty (Baker et al., 2016) and convicts' sentences (Wilson & Rule, 2015, 2016).

After controlling for men's histories of delinquency and substance use, their parents' income during childhood negatively related to the men's adult facial trustworthiness. A

disadvantaged childhood may therefore manifest in later facial characteristics, potentially leading to further disadvantages in legal contexts and healthcare.

It is worth highlighting that the outcomes associated with tobacco use in the study were measured in early adulthood. Whereas many smoking-related negative outcomes typically emerge later in life (e.g., lung cancer and cardiovascular disease; Department of Health and Human Services, 1994; Gavin, 2004), the present findings indicate that tobacco use relates to facial trustworthiness (and attractiveness) much earlier. If these associations reflect causal and modifiable processes , then these findings may have important implications for prevention. Namely, knowledge of the short-term negative consequences of substance use on personal appearance may particularly alarm young people, and could prove more salient than warnings based on long-term health risks. Indeed, reminders about the impact of tobacco use on facial wrinkling increases motivation to quit smoking among young women (Weiss et al., 2010).

Because nearly 90% of adult smokers started using tobacco by age 18 (Johnston et al., 2018), education about these additional negative impacts on facial appearance could further discourage onset and encourage cessation among young people. The present study suggests that these negative impacts go beyond premature aging and attractiveness. Moreover, given that facial untrustworthiness may affect clinicians' decisions when delivering medical treatment (Mattarozzi et al., 2017), its influence may particularly apply to chronic smokers, who incur more costs from inpatient and outpatient care than nonsmokers do (with such expenditures providing a proxy for more frequent or more serious need; Sturm, 2002).

The present study had multiple design strengths, including strong retention, the ecological validity of the photographs, use of multiple measurement methods that minimize numerous confounds and biases (i.e., self-reports, arrest records, and strangers' ratings), and the

use of strangers' ratings that replicated prior research showing a great deal of agreement among raters as to who appears trustworthy versus untrustworthy (e.g., Todorov et al., 2009). Additionally, the participants' histories of criminal behavior and substance use render the sample highly relevant to research on facial trustworthiness and its implications for judicial proceedings. A longitudinal design is another study strength: following boys over more than a decade of their lives allowed measurement of early facial trustworthiness and the cumulative impacts of delinquency, tobacco use, and related variables.

Yet, the present study also had several limitations. First, given that the sample was of primarily White boys from at-risk neighborhoods in the 1980s, the findings may not generalize to contemporary time periods, women, and other ethnic groups; sample characteristics also precluded examination of racial bias. Second, the measurement design (i.e., use of lifetime sums and means) had statistical advantages but also limits conclusions about the temporal sequence of predictors (e.g., delinquency and tobacco use). Third, we could not answer questions about how the frequency and timing of tobacco exposure may interact to predict facial trustworthiness (e.g., is early onset of moderate use worse than late onset of heavy use?) because tobacco use following onset typically escalates quickly to regular use and remains stable over time (Kerr et al., 2011; Lando et al., 1999). Future research should consider these questions, given that other health consequences of tobacco use increase with duration and amount, both of which increase over the lifespan when tobacco use begins earlier (Department of Health and Human Services, 2012).

Finally, the naturalistic nature of the photographs also imposed some methodological limits. Although the photographs were coded for early appearance and two primary influences on facial trustworthiness (i.e., positive affect and attractiveness), other characteristics might have

also contributed to strangers' perceptions (Hellstrom & Tekle, 1994; Nurmoja & Bachmann, 2014). Thus, additional unmeasured variables may play a role in the association between tobacco use and changes in facial trustworthiness. Whereas these limitations reduced our ability to identify the mechanisms underlying the observed associations, the naturalistic quality of the photographs bolsters the external validity of the study; that is, it supports the idea that signs of tobacco use may be detected from facial appearance by laypersons in the general environment.

Moreover, the observation that strangers' perceptions of men's current (but not earlier) appearances relates to their prior behavior is remarkable. Now that this association has been identified in temporal sequence (i.e., tobacco use predates reductions in facial trustworthiness), future research should replicate and extend these findings. For example, investigating the facial features that change with tobacco use over time (e.g., curve of the lips) may help to illuminate the causal or circumstantial nature of these associations.

Conclusion

Adult (but not early adolescent) facial trustworthiness relates to problem behavior histories. Here, prospective reports of tobacco use across adolescence and early adulthood predicted strangers' perceptions of adult facial trustworthiness after controlling for other forms of substance use, delinquency, family income, and facial trustworthiness in early adolescence. This study opens new lines of inquiry into social bias, facial characteristics important to person perception, and potential consequences of tobacco use beyond those linked to physical health. As such, the findings may bear utility for novel applications, such as tobacco-use prevention.

References

Adams, R. B., Jr., Garrido, C. O., Albohn, D. N., Hess, U., & Kleck, R. E. (2016). What facial appearance reveals over time: When perceived expressions in neutral faces reveal stable emotion dispositions. Frontiers in Psychology, 7, 986.

Ambady, N, & Skowronski J. J. (2008). *First impressions.* New York, New York: The Guilford Press.

Ary, D. V., Duncan, T. E., Biglan, A., Metzler, C. W., Noell, J. W., & Smolkowski, K. (1999). Development of adolescent problem behavior. *Journal of Abnormal Child Psychology, 27*, 141-150.

Baker, A. Porter, S., ten Brinke, L., & Mundy, C. (2016). Seeing is believing: Observer perceptions of trait trustworthiness predict perceptions of honesty in high-stakes emotional appeals. *Psychology, Crime & Law, 22,* 817-31.

Bjornsdottir, R. T., & Rule, N. O. (2017). The visibility of social class from facial cues. *Journal of Personality and Social Psychology, 113*, 530-546. doi: 10.1037/pspa0000091

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2013). The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General, 142*(1), 143.

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2015). Face-ism and kernels of truth in facial inferences. *Trends in Cognitive Sciences, 19*(8), 421-422.

Capaldi, D. M., Chamberlain, P., Fetrow, R. A., & Wilson, J. E. (1997). Conducting ecologically valid prevention research: Recruiting and retaining a "whole village" in multimethod, multiagent studies. American Journal of Community Psychology, 25, 471-492.

Capaldi, D. M. & Patterson, G. R.. (1989). *Psychometric properties of fourteen latent constructs from the Oregon Youth Study.* Springer-Verlag: New York.

Ciapponi, A. (2014). *Systematic review of the link between tobacco and poverty.* Geneva,

Switzerland: World Health Organization. Retrieved from:

http://apps.who.int/iris/bitstream/10665/136001/1/9789241507820_eng.pdf

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from

faces: A developmental study. *Psychological Science*, 25, 1132-1139. doi:

10.1177/0956797614523297

Cohen, S., Janicki-Deverts, D., Chen, E., & Matthews, K. A. (2010). Childhood socioeconomic

status and adult health. *Annals of the New York Academy of Sciences, 1186*, 37-55.

Collins, M. A., & Zebrowitz, L. A. (1995). The contributions of appearance to occupational

outcomes in civilian and military settings. *Journal of Applied Social Psychology, 25,* 129-

163.10.1111/j.1559-1816.1995.tb01588.x

Department of Health and Human Services USA. (1994.) *Preventing tobacco use among young

people: A report of the Surgeon General.* US Department of Health and Human Services.

Dion, K. K. (1973). Young children's stereotyping of facial attractiveness. *Developmental

Psychology, 9*, 183-188.

Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: the role of appearance in peer-to-

peer lending. *The Review of Financial Studies*, *25*, 2455-2484.

Elliott, D.S. Ageton, S.S., Huizinga, D., Knowles, B.A., & Canter, R.J. (1983). The prevalence

and incidence of delinquent behavior: 1976–1980. National estimates of delinquent

behavior by sex, race, social class, and other selected variables. In *National youth survey

report No. 26.* Boulder, CO: Behavioral Research Institute.

Freiman, A., Bird, G., Metelitsa, A.I., Barankin B., Lauzon, G.J. (2004.) Cutaneous effects of

smoking. *Journal of Cutaneous Medicine and Surgery, 8*, 415-23.

Gavin, A. (2004). Smoking is a major cause of premature death worldwide. *Evidence-based Healthcare, 8*, 95-96.

Gilbert, D. T. (1995). Attribution and interpersonal perception. In Tesser, A (Ed.), *Advanced Social Psychology.* (99-147). Boston, Massachusetts: McGraw-Hill.

Hellström, Å., & Tekle, J. (1994). Person perception through facial photographs: Effects of glasses, hair, and beard on judgments of occupation and personal qualities. *European Journal of Social Psychology, 24*(6), 693-705.

Jessor, R. (1991). Risk behavior in adolescence: a psychosocial framework for understanding and action. *Journal of adolescent Health, 12*(8), 597-685.

Johnston, L. D., Miech, R. A., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E. & Patrick, M.E. (2018). *Monitoring the Future national survey results on drug use, 1975-2017: Overview, key findings on adolescent drug use.* Ann Arbor: Institute for Social Research, The University of Michigan. Retrieved from http://www.monitoringthefuture.org/pubs/monographs/mtf-overview2017.pdf

Kerr, D. C .R., Owen, L. D., & Capaldi, D .M. (2011). The timing of tobacco use onset, prolonged abstinence, and relapse in men: A prospective study from ages 18 to 32 years. *Addiction, 106*, 2031-2038.

Kerr, D. C. R., Capaldi, D. M., Owen, L. D., Wiesner, M., & Pears, K. C. (2012). Change in at-risk American men's crime and substance use trajectories following fatherhood. *J Marriage Fam., 73*(5), 1101-1116.

Kleiman, S., & Rule, N.O. (2013). *Detecting suicidality from facial appearance. Social Psychological and Personality Science, 4,* 453-60

Lando, H. A., Thai, D. T., Murray, D. M., Robinson, L. A., Jeffery, R. W., Sherwood, N. E., & Hennrikus, D. J. (1999). Age of initiation, smoking patterns, and risk in a population of working adults. *Preventative Medicine, 29,* 590-598.

Ma, F., Xu, F., & Luo, X. (2015). Children's and adults' judgments of facial trustworthiness: the relationship to facial attractiveness. *Perceptual and Motor Skills, 121*(1), 179-198.

Ma, F., Xu, F., & Luo, X. (2016). Children's facial trustworthiness judgments: Agreement and relationship with facial attractiveness. *Frontiers in Psychology, 7*, 499.

Malatesta, C. Z., Fiore, M. J., & Messina, J. J. (1987). Affect, personality, and facial expressive characteristics of older people. *Psychology and Aging*, *2*, 64-69. doi: 10.1037/0882-7974.2.1.64

Mattarozzi, K., Colonnello, V., De Gioia, F., & Todorov, A. (2017). I care, even after the first impression: Facial appearance-based evaluations in healthcare context. *Social Science & Medicine*, *182*, 68-72.

Mattarozzi, K., Todorov, A., Marzocchi, M., Vicari, A., & Russo, P. M. (2015). Effects of gender and personality on first impressions. *PLoS one*, *10*(9), e0135529.

Muthén, L. K., & Muthén, B. O. (1998-2011). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Morita, A. (2007). Tobacco smoke causes premature skin aging. *Journal of Dermatological Science, 48*(3), 169-175.

Nurmoja, M., & Bachmann, T. (2014). Consistency of the relation between individual facial cues and the perceived suggestibility and trustworthiness across varying facial expressions. *TRAMES, 18*(68/63), 2, 105-119.

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., ... & Caspi, A. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and psychopathology, 20*(2), 673-716.

Oosterhof, N. N. & Todorov, A. (2008). The functional basis of face evaluation. *PNAS, 105*(32), 11087-11092.

Okada, H. C., Alleyne, B., Varghai, K., Kinder, K., Guyuron, B. (2013). Facial changes caused by smoking: A comparison between smoking and nonsmoking identical twins. *Plastic and Reconstructive Surgery, 132*(5), 1085-1092.

Re, D. E., & Rule, N. O. (2015). Appearance and physiognomy. In D. Matsumoto, H. Hwang, & M. Frank (Eds.), *APA Handbook of Nonverbal Communication* (pp. 221-256). Washington, DC: American Psychological Association. doi:10.1037/14669-009

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology, 104,* 409–426

Schäfer, T., Nienhaus, A., Vieluf, D., Berger, J., & Ring, J. (2001). Epidemiology of acne in the general population: the risk of smoking. *British journal of dermatology, 145*, 100-104.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47,* 609-612. doi: 10.1016/j.jrp.2013.05.009

Skinner, A. L. Woods, A., Stone, C. J., Penton-Voak, I., & Munafo, M. R. (2017). Smoking status and attractiveness among exemplar and prototypical identical twins discordant for smoking. *Royal Society Open Science, 4* (161076). *http://dx.doi.org/10.1098/rsos.161076*

Slepian, M. L., & Ames, D. R. (2015). Internalized impressions: The link between apparent facial trustworthiness and deceptive behavior is mediated by targets' expectations of how

they will be judged. *Psychological Science, 27*(2), 282-288. DOI:

10.1177/0956797615594897

Sofer, C., Dotsch, R., Wigboldus, D. H., & Todorov, A. (2015). What is typical is good: The

influence of face typicality on perceived trustworthiness. *Psychological Science, 26*(1),

39-47.

Stillman, T. F., Maner, J. K., & Baumeister, R. F. (2010). A thin slice of violence:

Distinguishing violent from nonviolent sex offenders at a glance. *Evolution &*

Human Behavior, 31, 298–303.

Stirrat, M. & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width

and trustworthiness. *Psychological Science, 21*(3), 349-54.

Sturm R. (2002.) The effects of obesity, smoking, and drinking on medical problems and costs.

Health Aff. 21(2), 245-53.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from

faces: Determinants, consequences, accuracy, and functional significance. *Annual Review*

of Psychology, *66*, 519-545.

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model

based approach. *Social Cognitive and Affective Neuroscience, 3*(2), 119-127.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after

minimal time exposure. Social Cognition, 27, 813–833. doi:10.1521/soco.2009.27.6.813

U.S. Public Health Service. Office of the Surgeon General (2012). Preventing Tobacco Use

Among Youth and Young Adults: A Report of the Surgeon General. Atlanta (GA):

Centers for Disease Control and Prevention (US); National Center for Chronic Disease

Prevention and Health Promotion (US) Office on Smoking and Health.

Van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*, 796-803.

Watts, W. D., & Wright, L. S. (1990). The relationship of alcohol, tobacco, marijuana. *Adolescence, 25*(97), 171.

Webb, N. M., Shavelson, R. J., & Haertel, E. H. (2006). Reliability coefficients and generalizability theory. *Handbook of Statistics*, *26*, 81-124. doi: 10.1016/S0169-7161(06)26004-8

Weiss, C., Hanebuth, D., Coda, P., Dratva, J., Heintz, M., & Stutz, E. Z. (2010). Aging images as a motivational trigger for smoking cessation in young women. *Int J Environ Res Public Health, 7*(9), 3499-3512.

Wiesner, M., Capaldi, D. M., & Kim, H. K. (2007). Arrest trajectories across a 17-year span for young men: Relation to dual taxonomies and self-reported offense trajectories. *Criminology*, *45*(4), 835-863.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological science, 17*(7), 592-598.

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, *26*, 1325-1331.

Wilson, J. P., & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes: The role of facial trustworthiness. *Social Psychological and Personality Science*, *7*, 331-338.

Wilson, J. P., & Rule, N. O. (2017). Advances in understanding the detectability of trustworthiness from the face: Toward a taxonomy of a multifaceted construct. *Current Directions in Psychological Science, 26*, 396-400.

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Boulder, CO: Westview Press.

Zebrowitz, L. A., Collins, M. A., & Dutta, R. (1998). Appearance and personality across the

lifespan. *Personality and Social Psychology Bulletin*, *24*, 736-749.

Zebrowitz, L. A., & Montepare, J. M. (1992). Impressions of Babyfaced Individuals Across the

Life Span. *Developmental Psychology, 28,* 1143-1152. doi:10.1037/0012-1649.28.6.1143

Zebrowitz, L. A. & Montepare, J. M. (2008). First impressions from facial appearance cues. In

N. Ambady & J. J. Skowronski (Eds.), *First impressions* (171 - 204). New York, New

York: The Guilford Press.

Zebrowitz, L. A., Olson, K., & Hoffman, K. (1993). Stability of babyfaceness and attractiveness

across the life span. *Journal of Personality and Social Psychology, 64,* 453-466.

doi:10.1037/0022-3514.64.3.453

PROBLEM BEHAVIOR AND FACIAL TRUSTWORTHINESS

Table 1

Descriptive Statistics and Correlations Among Study Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age-24 trustworthiness	-	.52**	.40**	-.39**	-.19*	-.08	-.19*	-.18*	-.19*	-.23**	-.26**	.27**	.28**
2. Age-24 attractiveness		-	.06	-.20**	.05	-.01	-.03	.00	-.03	-.12	-.08	.12	.13
3. Age-24 positive affect			-	-.11	-.06	.01	.01	-.01	-.05	-.11	-.10	.11	.10
4. Tobacco use (thousands)				-	.34**	.36**	.38**	.40**	.35**	.37**	.47**	-.19*	-.08
5. Marijuana use					-	.53**	.66**	.87**	.69**	.46**	.66**	-.17*	-.02
6. Alcohol use						-	.37**	.75**	.62**	.31**	.59**	-.02	.08
7. Other-drug use							-	.81**	.61**	.38**	.58**	-.14	-.01
8. Other substance use†								-	.79**	.47**	.75**	-.13	.02
9. Delinquency									-	.46**	.86**	-.20**	.03
10. Arrests										-	.79**	-.28**	-.01
11. Delinquency/arrests											-	-.31**	.01
12. Parent income												-	.12
13. Age-14 trustworthiness													-
Mean	3.76	2.12	1.94	3.43	1.05	2.50	0.26	0.00	0.00	4.32	0.00	0.00	3.81
Standard deviation	0.59	0.55	1.06	3.54	1.28	1.04	0.48	1.00	1.00	6.91	1.00	0.93	0.66

PROBLEM BEHAVIOR AND FACIAL TRUSTWORTHINESS

Minimum		2.40	1.19	1.00	0.00	0.00	0.00	0.00	-1.91	-1.40	0.00	-1.38	-1.44	2.00
Maximum		5.20	4.00	5.00	10.99	5.04	4.75	3.40	3.76	2.88	39.00	2.54	1.98	5.53
N		177	177	177	177	177	177	177	177	177	177	177	177	159

Notes. * $p < .05$, ** $p < .01$.

†Composite of non-tobacco (alcohol, marijuana, and other drug) substance use.

Table 2

Regression Analysis Predicting Age-24 Facial Trustworthiness (n = 177)

Predictor	Age	b (SE)	β [95% CI]	p
Parent income	10-12	.10 (.04)	.14 [.01, .26]	.03
Age-14 trustworthiness	13	.18 (.06)	.21 [.08, .34]	.001
Tobacco use	12-23	-.05 (.01)	-.29 [-.42, -.15]	<.001
Delinquency/arrests	13-23	-.02 (.06)	-.03 [-.24, .17]	.75
Other substance use	12-23	-.02 (.06)	-.03 [-.22, .16]	.76
Positive affect	24	.18 (.04)	.33 [.21, .46]	<.001

Note. b (SE) = parameter estimate (standard error). Significant covariances ($p < .05$) emerged for delinquency and parent income, *cov*(*SE*) = -.31(.07); for tobacco use with both parent income, *cov*(*SE*) = -.18(.07), and delinquency, *cov*(*SE*) = .47(.06); and for other drug use with both delinquency, *cov*(*SE*) = .75(.03), and tobacco use, *cov*(*SE*) = .40(.06). The covariance between age-14 trustworthiness and tobacco use was not significant, *cov*(*SE*) = -.089(.08), *p* = .27.