Core Self-Evaluations Mediate the Association between Leaders' Facial Appearance and

Their Professional Success: Adults' and Children's Perceptions

Erik Dietl

Business and Organizational Psychology,

University of Hohenheim, Germany

Nicholas Rule

Department of Psychology,

University of Toronto, Canada

Gerhard Blickle

Industrial, Organizational, and Economic Psychology,

University of Bonn, Germany

Abstract

Although the link between facial appearance and success is well established, the mechanisms responsible for this association have remained elusive. Evolutionary theory suggests that perceived leadership characteristics should be important for men's self-concept. Drawing on implicit leadership theory and evolutionary perspectives, we therefore examined the associations between first impressions based on facial appearance, core self-evaluations (CSEs), leadership role occupancy, and career success among a sample of working men. In Study 1, we found that CSEs mediated the relationship between individuals' facial appearance and measures of their success as leaders. In Study 2, we replicated these results using children's ratings of facial appearance, thus suggesting that basic properties of the targets' faces communicated their leadership ability more than the perceivers' life experience or acquired knowledge. These results suggest that people may use facial appearance as a diagnostic tool to determine the leadership ability of others.

Keywords: career success, core self-evaluations, evolution, face perception, leadership role occupancy

Core Self-Evaluations Mediate the Association between Leaders' Facial Appearance and

Their Professional Success: Adults' and Children's Perceptions

Parents commonly admonish their children not to judge others based on their

appearance but, rather, to look "inside" a person to his or her personality, intelligence, and

values. In many cases, this may be sound advice. Yet, in other instances, outside appearance

may honestly convey some of what lies underneath. For example, people's basic personality

traits (Penton-Voak, Pound, Little, & Perrett, 2006), political opinions (Samochowiec,

Wanke, & Fiedler, 2010), intelligence (Zebrowitz, Hall, Murphy, & Rhodes, 2002), and

leadership success (Rule & Ambady, 2008) all correlate, albeit weakly, with subjective

assessments derived from their facial appearance (cf. Antonakis & Eubanks, 2017). In

addition, facial appearance can influence one's career and leadership success in various

domains (e.g., Antonakis, 2011; Hosoda et al., 2003). The mechanisms responsible for these

associations have remained elusive, however.

Here, we sought to more directly map out how leadership relates to facial appearance

by testing the link between one's outer appearance and inner self-concept, hypothesizing that

individuals' self-concepts would mediate the association between their appearance and

leadership success. In two studies, we investigated how self-concept mediates the positive

association between both children's and adults' subjective assessments of appearance with

working men's professional success, as measured by their leadership role occupancy and

occupational status. Doing so produced three key contributions.

First, we answered Baruch and Bazionelus's (2011) recent call for research on the

mechanisms linking job-irrelevant human capital and success outcomes by examining core

self-evaluations (CSEs; the "fundamental assessments that people make about their

worthiness, competence, and capabilities;" Judge, Bono, Erez, & Locke, 2005, p. 257) as a

mediator of the positive association between facial appearance and leader's professional

success. This allowed us to extend prior work showing that CSEs link individuals'

attractiveness to their income by examining how facial appearance relates to leadership

attainment via CSEs (Judge, Hurst, & Simon, 2009).

Second, by sampling judgments of leadership from both adults (Study 1) and children

(Study 2), we tested and extended theories suggesting that individuals without prior

experience can reliably intuit others' leadership ability with the face serving as a diagnostic

tool (Spisak, Dekker, Kruger, & Van Vugt, 2012). Concordant results from adults and

children would suggest that first impressions based on facial appearance may not completely

arise from socialized experience within the work environment but may also stem from a

possibly innate sensitivity to detect hierarchical cues. Moreover, if leadership appearance

relates to individuals' CSEs, selecting these leaders might suggest a "kernel of truth" in

leadership judgments by reflecting internal traits associated with effective leadership and

work behaviors (Berry, 1990; Chang, Ferris, Johnson, Rosen, & Tan, 2012; Hu, Wang, Liden,

& Sun, 2012; Penton-Voak et al., 2006). In addition, by examining how both adults' and

children's leadership judgments relate to leadership attainment, we attempted to conceptually

replicate previous work on political leaders (Antonakis & Dalgas, 2009), thereby providing a

much needed replication of an intriguing finding in another context (i.e., the world of work;

see Antonakis, 2017).

Third, although evidence suggests that facial appearance predicts success among

political, military, and business leaders, this research has focused almost exclusively on elite

leaders (see Re & Rule, 2015, for review). Thus, because empirical studies examining the

association between facial appearance and leadership success in the world of work are scarce

(limiting the generalizability of prior studies), we tested how facial appearance and leadership

outcomes relate among nonelite leaders across a broad variety of jobs and industries, thereby

allowing for generalization beyond a specific company of profession. In addition, our study

heeds the call that organizational and management research should also focus on "lower-

echelon employees" (Bamberger & Pratt, 2010), who represent the majority of people working in organizations.

Facial Appearance and Leadership

Several recent studies have shown that people can reliably infer leaders' success from mere photos of their faces. Judgments of personality traits (e.g., competence) predict the outcomes of elections in countries on almost every continent (Lawson, Lenz, Baker, & Myers, 2010; Little, Burriss, Jones, & Roberts, 2007; Martin, 1978; Rule et al., 2010; Todorov, Mandisodza, Goren, & Hall, 2005) and correlate with the amounts of profits that chief executive officers (CEOs) earn for their companies (Harms, Han, & Chen, 2012; Rule & Tskhay, 2014). Despite the importance of leadership evaluations (such as deciding for whom to vote), quick and unreflective judgments predict these outcomes better than more thoughtful assessments do (Ballew & Todorov, 2007). More astonishing, even children's subjective judgments of political candidates predict electoral outcomes just as effectively as adults' evaluations do (Antonakis & Dalgas, 2009).

Other studies have suggested that these superficial judgments may probe individuals' actual dispositions. For example, Mueller and Mazur (1996) found that perceptions of West Point cadets' facial dominance predicted their military ranks at the ends of their careers. Similarly, perceptions of corporate lawyers' personality traits in college predicted their accomplishments as leaders of their firms as much as 50 years later (Rule & Ambady, 2011).

Implicit leadership theory (ILT; Lord, Foti, & De Vader, 1984) may help to explain why facial appearance relates to leadership success outcomes. The theory explains how leaders emerge, why someone is accepted as a leader, and why a leader can exert influence upon others. Individuals' ILTs refer to representations nonconsciously held by followers that help discriminate "leaders" from "non-leaders" and facilitate instant assessments of the leadership qualities of another person (Shondrick & Lord, 2010). Observers use these prototypes to automatically determine whether a leader matches their prototypical

expectations. Individuals who match the leader prototype are considered more favorably and can emerge more easily as leaders (Lord et al., 1984). Consistent with these arguments, research has shown that first impressions of leadership from faces were higher when facial appearance and expressions matched the observers' prototypes (Trichas & Schyns, 2012). Furthermore, ILTs guide judgments of leadership from nonverbal cues, which relate to measures of leaders' success (Tskhay, Xu, & Rule, 2014).

People with good ideas frequently struggle to implement them because they cannot inspire others to follow them or adopt their ideas. Leadership may thus require the ability to entrain other people towards compliance, commitment, and positive affect (Pfeffer, 1981). Yet, despite these important functions, individuals' height, sex, or facial appearance may influence whether others view them as leaders (e.g., Elgar, 2016; Re et al., 2013). Thus, prototypical images of a leader, rather than substantive evaluations of their skills, may shape followers' perceptions of whether someone is worth following. Followers' and observers' perceptions of faces and their attributions of leadership qualities are influenced by ILTs as well (e.g., Antonakis, 2011; Trichas & Schyns, 2012). Therefore, favorable leader perceptions based on one's face may relate to an individual's leadership success. Taken together, perceptions of leadership ability from the face (whether through direct assessments of how effective a leader looks, or indirectly through judgments of traits like competence and dominance) correlate with elite leaders' success. In an attempt to replicate prior work, we therefore expected:

Hypothesis 1. Perceptions of leadership from the face positively relate to leadership role occupancy and career success for people working at various levels of leadership within an organization.

Facial Appearance, Core Self-Evaluations, and Success

Job-irrelevant human capital denotes individual characteristics that logically should not relate to job performance but that nonetheless seem to influence career success (Baruch &

Bozionelos, 2011). Facial appearance is a typical example of job-irrelevant human capital that

nevertheless relates to career success and leadership emergence (Antonakis & Eubanks, 2017;

Hosoda et al, 2003). Hence, understanding the mechanisms that mediate the association

between facial appearance and workplace outcomes can expand models seeking to predict

organizational efficiency and productivity.

The mechanisms responsible for these associations have remained elusive, however.

We contend that favorable perceptions of a face may not only relate to an individual's

leadership success, but may also influence that person's self-concept. For instance, an

attractive appearance can elicit positive expectations and stereotypes (Dion, Berscheid, &

Walster, 1972), and these positive impressions can turn into self-fulfilling prophecies

whereby one gradually adopts the traits and behaviors that others expect, allowing the person

to develop a favorable self-concept (e.g., higher self-esteem, self-efficacy; Antonakis, 2011;

Langlois et al., 2000; Rosenthal & Jacobsen, 1968; Snyder, Tanke, & Berscheid, 1977).

Indeed, Judge, Hurst, and Simon (2009) found that CSEs mediated the association between

individuals' attractiveness and their income. In this vein, we investigated whether CSEs might

similarly mediate the relationship between perceptions of leadership from facial appearance

and two measures of leadership success: leadership role occupancy and occupational status.

Though often considered a stable trait, experiences can also influence CSEs (Nübold,

Muck, & Maier, 2013; Wu & Griffin, 2012). Extending and building on prior work on facial

appearance and self-concept, we thus propose that perceptions of leadership from the face

(e.g., competence, trustworthiness) should positively influence individuals' global self-

concept in three ways. First, observers (e.g., coworkers, clients, supervisors, mentors) might

treat individuals as if leader-like facial features reliably indicate their true underlying skills

and, consequently, trust them more (Rezlescu, Duchaine, Olivola, & Chater, 2012). This

positive treatment is likely to have a strong influence on global self-worth (Harter, 2006).

Second, individuals who look leader-like might be treated and accepted as leaders more often,

positively affecting their CSEs and providing them more opportunities to develop their

(leadership) skills—in other words, a self-fulfilling prophecy (e.g., Antonakis, 2011;

Friedman & Zebrowitz, 1992; Rule & Ambady, 2011). Consistent with these ideas,

individuals with the right look may be more confident in their abilities and more likely to be

considered leaders (DeRue & Ashford, 2010; Judge et al., 2009; Mobius & Rosenblat, 2006).

Third, having a facial appearance that garners impressions of high status and leadership

should correlate with a positive global self-concept, especially among men. Indeed,

evolutionary models of mate selection suggest that men must advertise their status and

resources to succeed in mating because women seek these qualities in potential mates (Buss,

1989; Li et al., 2002). Accordingly, Campbell and Wilbur (2009) found that status-related

attributes (e.g., having a good job, being financially secure, seeking status) importantly

influenced men's self-concepts. These qualities should therefore contribute to the

development of higher CSEs when a man's face signals status and resources, as through the

appearance of leadership, which should prompt reactions from others that positively impact

their CSEs (Betzig, 1993; Buss, 2005). Taken together, we therefore hypothesized:

> *Hypothesis 2*. Perceptions of leadership from the face positively relate to CSEs.

Moreover, we expected that CSEs would mediate the positive association between

perceptions of leadership and career success. Along these lines, Chang et al.'s (2012) meta-

analysis showed that people with higher CSEs performed better at their jobs and enacted

fewer harmful behaviors against their organizations and fellow employees. They further found

that CSEs positively correlated with individuals' income, suggesting that greater CSEs may

encourage occupational status and higher leadership role occupancy (see also Judge, Bono,

Ilies, & Gerhardt, 2002; Li, Arvey, & Song, 2011). Men perceived as better leaders may

therefore achieve greater career success (e.g., occupational status) and better leadership roles

because they develop a more positive global self-concept (i.e., have higher self-esteem, self-

efficacy, a stronger internal locus of control, and emotional stability). We therefore

hypothesized:

> *Hypothesis 3.* CSEs mediate the positive association between perceptions of
>
> leadership from the face and both (a) leadership role occupancy and (b) career
>
> success.

Face Perception and Evolutionary mechanisms.

General adaptive mechanisms of face perception, and more specific mechanisms of

leadership perception, suggest that humans may have innate mechanisms to quickly recognize

others' leadership ability (Short et al., 2012; Spisak, Dekker et al., 2012). Because both

children and "naïve" undergraduates can reliably predict meaningful leadership outcomes, it

appears that individuals without leadership experience in an employment setting can reliably

intuit others' leadership ability and that the face serves as a diagnostic tool in these judgments

(e.g., Antonakis & Dalgas, 2009). To test this, we sampled leadership judgments from both

adults (Study 1) and children (Study 2) to investigate whether individuals' ability to detect

leadership success from leaders' faces arises from the experience of working with leaders

(i.e., the correlation emerges only for adults' judgments) or from a more basic sensitivity to

cues that signal leadership (i.e., the correlation emerges for both adults' and children's

judgments). Moreover, given that CSEs reflect valid leadership abilities and relate to effective

work behaviors (Chang et al., 2012; Hiller & Hambrick, 2005; Hu et al., 2012; Judge, Piccolo,

& Kosalka, 2009), establishing that CSEs mediate the relation between perceived leadership

judgments and actual leadership success may suggest that perceivers reliably detect kernels of

truth about leaders' abilities from their faces.

Several studies have suggested that people evolved acuity in interpersonal perception

to help them resolve recurrent social problems, such as detecting aggression and other related

threats to mating and survival (Sell et al., 2009; Short et al., 2012; Spisak, Dekker et al.,

2012). Due to their critical survival-enhancing contributions, these mechanisms may function

independently of prior experience (e.g., Spisak, Dekker et al., 2012). For example, estimates of aggression made by untrained adults and 8-year-old children correlated with aspects of male facial structure related to actual aggressive behavior (Short et al., 2012). People may possess similar mechanisms for perceiving leadership. Van Vugt, Hogan, and Kaiser (2008) have suggested that leadership might have evolved as a strategy for resolving problems such as group movement, intragroup conflict, and intergroup competition. Successes and failures of leadership might therefore create enough variation between groups that natural selection for leadership could operate at the group level under certain conditions (e.g., if resources were scarce and well-led groups performed better at hunting, food-sharing, and warfare; Wilson, Van Vugt, & O'Gorman, 2008). Faster and more effective group coordination can be essential in emergencies and lead to a first-mover advantage when exploiting resources or attacking other groups (Van Vugt et al., 2008). Identifying individuals with the competence and expertise for leadership can thus have high value, promoting cognitive mechanisms for doing so (see Lord et al., 1984; Spisak, Dekker et al., 2012; Spisak, Homan, Grabo, & Van Vugt, 2012).

Leadership categorization theory thus posits that individuals resembling the leader prototype are more likely to emerge as leaders (Lord et al., 1984), to which Spisak, Homan et al. (2012) added the suggestion that leader prototypes evolved to include both psychological (e.g., socialized and culturally-specific ILTs) and physical markers (e.g., the face). Categorizing leaders versus nonleaders from their faces could therefore facilitate leader emergence across situations and in specific contexts (e.g., Boggild & Laustsen, 2016; Laustsen & Petersen, 2017; Little et al., 2007); for instance, masculine and feminine facial cues facilitate leader emergence during intergroup conflict and cooperation, respectively (e.g., Spisak, Homan et al., 2012).

Current Work

We tested our three hypotheses across two cross-sectional field studies. In Study 1, we investigated whether CSEs mediated the anticipated positive associations between individuals' facial appearance and two measures of their success as leaders (leadership role occupancy and career success) via direct assessments of leader choices based on the face. In addition, we collected trait judgments (e.g., trustworthiness, dominance) to explain adults' leadership preferences. In Study 2, we extended and generalized this mediation model by replicating it with a sample of child participants. Overall, we sought to measure whether perceptions of leadership predict the success of nonelite leaders from a working population and to test the role that CSEs might play in linking perceptions of leadership ability to actual leadership success. Studying these judgments within an unconventional and working population should permit greater generalization of leadership theory to diverse settings. Moreover, corroborating such an association with children could help to buttress the potentially inherent basis for the cognitive-perceptual mechanisms supporting this link, thereby helping to establish the fundamental nature of leadership perception among both targets and perceivers.

Study 1

Method

Stimuli. Target persons worked across a broad range of occupations and industries. They were at least 29 years old because leadership and career success need time to evolve in the world of work. All targets shared handball sports as a common interest and were similarly active handball players or coaches, allowing us to roughly equate physical fitness and its correlates in our target sample (see Ehrenspiel & Strahler, 2012).[1] As mentioned above, this unconventional context allows for generalization of leadership theory to a broader context

---

[1] Interested readers may also wish to consult Elgar's (2016) study of how physical stature relates to leader performance in Olympic and Paralympic teams.

than past studies, which have focused predominantly on elite leaders in high-status corporate and political positions.

We collected two head-shot photos of each of 130 Caucasian (M_{age} = 46.00 years, SD = 7.81) male players, who were either A-level coaches or players on senior teams (i.e., older than 29 years) within the German Handball Association. We only sampled people at least 29 years old because leadership and career success need time to evolve in the world of work and most adults have entered the labor force in Germany by 29 years of age (Weingerter, 2011). A-level coaches (who hold the top coaching license of the German Handball Association) were contacted through workshops, whereas senior players were contacted through their team coaches. The targets displayed a simple neutral expression in one of the photos. For the other photo, we instructed each target to present himself as a successful coach, taking as many as 10 photos and asking him to choose the one he thought best represented him as a leader by displaying the photos on the screen of the digital camera (henceforth referred to as the "impression-management" photo). We asked each person to take off his glasses and wear a cloak to obscure any clothing. Afterward, we standardized each photo's size so that the height from the chin to the top of the head was uniform across targets and converted each photo to grayscale to control for differences in lighting.

Measures.

Core self-evaluations. We measured targets' global self-concept with the German-validated version of the Core Self-Evaluations Scale (CSES; Judge et al., 2003; Stumpp et al., 2010). The CSES contains 12 self-report items with a response scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Sample items include *When I try, I generally succeed* and *Overall, I am satisfied with myself* (see Appendix A for all items). The internal consistency of the CSES was acceptable (α = .72).

Occupational status. Occupational status reflects the amount of power, prestige, and authority that society ascribes to a particular profession (e.g., Schooler & Schoenbach, 1994).

We asked targets to assess their occupational status from 1 (*unskilled worker*) to 15

(*proprietor of a large company*) using a scale frequently employed in sociology and

organizational behavior that is very similar to the Occupational Scale of Hollingshead's Index

of Social Position (see Blickle et al., 2011; Dietl, Meurs, & Blickle, 2017; Hartmann, 2002;

Miller & Salkind, 2002). Although self-reports of career success strongly correlate with

objective measures from archived company data (Judge, Cable, Boudreau, & Bretz, 1995), we

validated the association by manually coding a random subsample of 17 targets' reported

professions using the International Standard Classification of Occupations 2008 (ISCO-08;

Züll, 2016) to calculate two indicators of occupational status: the International Socio-

Economic Index (ISEI) of occupational status (Ganzeboom, 2010) and the Standard

Occupational Prestige Scale (SIOPS, Ganzeboom & Treiman, 2003; Treiman, 1977). Both

indices strongly correlated with the self-reported occupational status scale: $r(15) = .63$, $p =$

.006, and $r(15) = .70$, $p < .001$, respectively). Both the target's occupational status and

leadership role (described below) referred to full-time jobs that were *outside* the sports leisure

activity through which they were recruited, except for 10 targets who worked full-time as

professional coaches.

     Leader role occupancy. We measured the leadership role that each individual

occupied with a scale consisting of different hierarchical levels in organizational leadership.

Specifically, targets reported their management level from 1 (*no leadership position*) to 5

[*upper management level (executive board)*]. Previous research using this scale (Blickle,

Witzki, & Schneider, 2009) revealed positive associations between the respondents' scores

and their salary, $r(279) = .30$, $p < .001$, hierarchical position within their company (0% =

bottom, 100% = top), $r(300) = .51$, $p < .001$, and number of employees that they supervise,

$r(285) = .24$, $p < .001$.[2]

---

[2] Based on reanalysis of original published data; full results available from the first author.

Control variables. Because working as a professional coach might relate to CSEs, occupational status, and leadership, we modeled our data with and without controlling for whether each target worked full time as a professional coach using a dummy variable (0 = not a professional coach; 1 = professional coach; see Becker et al., 2015).[3]

Participants. In order to sample participants from a homogeneous context, we contacted active and former handball players via invitations to follow links to an online survey. Participants were recruited either personally at their sports clubs or via public postings in online social networks and forums frequently used by handball players. A total of 366 participants started the survey and 276 (194 male, 82 female; M_{age} = 29.73 years, SD = 11.33) completed the online questionnaire (completion rate 75.4%). Of these, 168 (60.8%) came from the working population, 71 (25.7%) were undergraduates, 33 (12.0%) were students at other levels, and four (1.4%) were occupied in some other way.

Procedure. We created an online survey using Inquisit 3.0.5.0 (2011). To keep the online study brief, we split the 260 photos into twelve sets following a 2 (photo type: neutral, impression-management) × 6 (target age: 29-39 years, 39-42 years, 43-46 years, 46-49 years, 49-54 years, and 54-65 years) between-subjects design with 22 photos in each set. Two targets (aged 39 and 46 years) appeared in two of the sets to yield equal numbers per set and because including them twice assured that we could accrue sufficient judgments from people who did not recognize them, as they were relatively famous handball coaches (indeed, we excluded 204 ratings of these targets for a total of 633 and 429 valid ratings). Moreover, the 46-49 year age set and 54-65 year age set overlapped with the 49-54 age set. In these cases, we randomly assigned three of the five 49 year-old targets and two of the four 54 year-old targets to one or the other group. Grouping the targets by age prevented age-related contrasts

---

[3] The targets completed a variety of measures in addition to those reported here; see Appendix B for a full list of all questions asked.

that might have affected participants' judgments (e.g., Krendl, Rule, & Ambady, 2014; see also Biernat & Manis, 1994).

We randomly assigned approximately 23 participants to rate one of the 12 sets of photos. The participants viewed all 22 faces in random order, blocked by rating type. Two of the blocks assessed leadership. In one, we asked the participants to "Imagine that you are playing on a capable team. Next season, your team could ascend to a higher league. You just need a good new coach. Several people introduce themselves. You get to co-decide who will be considered for preselection and could be your new coach. Should this person be considered for preselection?" and then rate each face on a 4-point scale from 1 (*Certainly not*) to 4 (*Definitely yes*). In the other, we asked them to "Imagine that you are working in a company. Because your department has been very successful, your supervisor has been promoted. The vacancy needs to be filled now. Several people introduce themselves. You get to co-decide who will be considered for preselection and could be your new supervisor. Should this person be considered for preselection?" using the same scale. The two leadership judgments correlated very strongly for both the neutral, $r(128) = .82$, $p < .01$, and impression-management photos, $r(128) = .80$, $p < .01$. We therefore averaged the ratings and refer to this composite variable as Choice as Leader.

Following this, the participants rated the faces along several trait dimensions (attractiveness, competence, dominance, facial maturity, likeability, strength of leadership, and trustworthiness) from 1 (*Not at all* X) to 7 (*Very* X) in random order. We selected these traits based on a multitude of previous studies implicating their importance for perceptions of leadership from faces (e.g., Rule & Ambady, 2008). Finally, the participants viewed each photo and indicated whether they recognized the person from outside the study. Consistent with prior research (e.g., Rule & Ambady, 2008), we were interested in naïve judgements and therefore removed the data for recognized faces prior to analysis (10.1% of all trials). In addition, we removed data from blocks in which the participant gave identical ratings to every

target (i.e., participants who did not follow the instructions and/or who finish the study

quickly), and excluded *in toto* data from participants who provided strings of identical ratings

in at least one third of the blocks (3.3% of all trials).[4]

Estimation strategy. We calculated single- and average-score intraclass correlation

coefficients (ICC1, ICC2; Bliese, 1998, 2002) to justify aggregating the judgments of the

targets for each rating.[5] We found a high degree of inter-rater agreement (all ICC2s > .80; see

Table 1). The judgments of the neutral and impression-management photos significantly

correlated for the Choice as Leader composite, $r(128) = .73$, $p < .01$, and all of the trait

ratings, $rs(128) \geq .57$, $ps < .01$; thus, we averaged the ratings for each judgment across the

two photo types (see the Supplemental Materials for results decomposed for the two photo

types separately).

We conducted a principal components analysis with varimax rotation for the

participants' trait ratings (i.e., competence, dominance, facial maturity, likeability, strength of

leadership, and trustworthiness). We did not include attractiveness because it would highly

correlate with the trait inferences and leadership ratings (due to common source and method

effects; Antonakis et al., 2010; Podsakoff et al., 2012), following previous research (Rule et

---

[4] The participants rated the likelihood that each person is an employee or leader, and the

likelihood that each person is a player or coach prior to the recognition question. We also

asked participants to self-report their education level, occupation, sports club, and sports

league alongside the other demographic variables before rating the targets. These variables

were exploratory and thus not included in the present analyses.

[5] Given that participants rated only some of the targets, we could not compute the indices

ICC(C,1) and *ICC(C,k)* – known as Cronbach's alpha – which control the inter-rater variance

and usually produce larger coefficients. For this reason, our reliability indices are probably

underestimated (McGraw & Wong, 1996).

al., 2010). Ratings of likeability and trustworthiness loaded together into a factor we named Warmth, and ratings of dominance and facial maturity loaded together into a factor we named Power (see Table 2 for factor loadings). Competence and strong leadership loaded highly on both Power and Warmth, however. Consistent with prior research (Rule et al., 2010), we therefore decided to remove competence and strong leadership from the exploratory analyses. We averaged the mean scores for each target into the composite variables: Power (dominance and facial maturity) and Warmth (likeability and trustworthiness).

Because we nested the targets' aggregated appearance ratings (Level-1) within photo sets (Level-2), we used a cluster-robust maximum likelihood (MLR) estimator with robust variance estimates to account for the nonindependence using Mplus 7.3 (Muthén & Muthén, 2012) and modeled the data using a mediation framework in which CSEs mediated the association between the Choice as Leader ratings and the outcomes (occupational status leadership role).

Although CSEs are often considered a trait, physical appearance and other variables (e.g., general mental ability and work experiences) may influence them (Judge et al., 2009; Wu & Griffin, 2012). One's occupational status or leadership role might even influence a person's CSE and self-efficacy (a component of CSEs considered state-like and malleable to mastery experience and social modeling; Bandura, 1986). The hypothesized mediation model may therefore be partly influenced by other omitted causes and the mediating variable (CSEs) is likely *endogenous* (Antonakis et al., 2010; Judge et al., 2009; Wu & Griffin, 2012), potentially leading to inconsistent parameter estimates. In addition, CSEs could also be endogenous because they share omitted common-method variance with the outcomes, which were rated by the same source (Antonakis et al., 2010).

To guard against endogeneity bias, we therefore used an instrumental-variable (IV) estimation method to obtain consistent parameter estimates for our potentially endogenous mediator variable (i.e., CSEs) using maximum likelihood estimation in Mplus 7.3 (Muthén &

Muthén, 2012; see also Antonakis et al., 2010; Shaver, 2005). We correlated the disturbance of the CSEs with the disturbances of the dependent variables (i.e., leadership role and occupational status) and regressed the outcomes on the mediator CSEs. Stable individual differences can serve as instruments so long as they are mostly exogenous (i.e., depend on genes and are not influenced by other variables) and no selection has taken place on them (Antonakis, 2011; Antonakis et al., 2010). We therefore used targets' age and Choice as Leader ratings to purge the CSEs of endogeneity bias (see Table 3 for descriptive statistics and correlations for all variables).

Results

Consistent with Hypothesis 1, cluster-robust regressions revealed that the targets' Choice as Leader scores positively related to their actual leadership role ($B = 0.72$, $SE = 0.27$, $p = .006$; Table 4, Model 3a) and occupational status ($B = 1.80$, $SE = 0.42$, $p < .001$; Table 4, Model 5a), regardless of whether the Professional Coach dummy variable was included (leadership role: $B = 0.67$, $SE = 0.27$, $p = .01$, Model 3b; occupational status: $B = 1.62$, $SE = 0.37$, $p < .001$, Model 5b). Moreover, Choice as Leader also positively related to CSEs ($B = 0.17$, $SE = 0.05$, $p = .001$; Table 4, Model 1a), regardless of whether the Professional Coach dummy variable was included ($B = 0.16$, $SE = 0.05$, $p = .002$; Table 4, Model 1b), supporting Hypothesis 2.

Instrumental-variable mediation analyses. Table 4 presents the summary of the path estimates for the IV-estimator regression models (Models 2a, 2b, 4a, 4b, 6a, 6b). CSEs significantly related to leadership role ($B = 4.61$, $SE = 2.01$, $p = .02$) and occupational status ($B = 10.57$, $SE = 4.35$, $p = .02$), regardless of whether the Professional Coach dummy variable was included (leadership role $B = 4.60$, $SE = 2.05$, $p = .03$; occupational status $B = 10.21$, $SE = 4.26$, $p = .02$). Cluster-robust mediation analyses testing Hypothesis 3 showed that CSEs mediated the positive associations between the Choice as Leader judgments and both success criteria: actual leadership role (indirect effect $= 0.74$, $SE = 0.18$, 95% CI [0.39, 1.10]; Table 4,

Models 2a, 4a) and occupational status (indirect effect = 1.71, *SE* = 0.32, 95% CI [1.08, 2.33]; Table 4, Models 2a, 6a), regardless of whether the Professional Coach dummy variable was included (leadership role indirect effect = 0.69, *SE* = 0.19, 95% CI [0.33, 1.06]; Table 4, Models 2b, 4b; occupational status indirect effect = 1.54, *SE* = 0.29, 95% CI [0.97, 2.10]; Table 4, Models 2b, 6b). In addition, both correlations between the disturbances of CSEs with leadership role (Ψ = −0.67, *SE* = 0.31, *p* = .03) and with occupational status (Ψ = −1.47, *SE* = 0.66, *p* = .03) were significant, regardless of whether the Professional Coach dummy variable was included. This indicates that CSEs are indeed endogenous and supports our use of instruments to obtain consistent estimates for the endogenous variable (Antonakis et al., 2010). Overidentification tests indicated that the mediation models fit the data well [$\chi^2(2)$ = 0.18, *p* = .92; with Professional Coach dummy: $\chi^2(2)$ = 0.17; *p* = .92]. The modification indices of the mediation models showed that none exceeded the minimum value, suggesting that our models fit the data well.

Exploratory analyses with trait inferences. The Choice as Leader judgments strongly associated with both Power, *r*(128) = .54, *p* < .001, and Warmth, *r*(128) = .90, *p* < .001. Cluster-robust regressions moreover showed that the targets' Warmth scores positively related to their actual leadership roles (*B* = 0.33, *SE* = 0.14, *p* = .02), occupational stata (*B* = 0.73, *SE* = 0.17, *p* < .001), and CSEs (*B* = 0.10, *SE* = 0.03, *p* = .004), whereas Power did not significantly relate to any of these (all /*B*/s < 0.57, all |*SE*|s < 0.50, all *p*s > .20). We again conducted instrumental variable mediation analyses, as described above, using the respective appearance ratings and targets' ages as instruments for their CSEs. CSEs mediated the positive associations between the Warmth judgments and both success criteria: actual leadership role (indirect effect = 0.41, *SE* = 0.11, 95% CI [0.19, 0.62]) and occupational status (indirect effect = 0.86, *SE* = 0.13, 95% CI [0.61, 1.12]). Yet, CSEs did not mediate the associations between Power and either success criterion (both |indirect effect|s < 0.27, |*SE*|s < 0.53, 95% CIs [-0.19, 0.42]).

Discussion

The results of Study 1 confirmed that people who look more like leaders have higher occupational status and leadership roles, and that targets' CSEs mediated the associations between perceptions of their leadership and these measures of their success. Thus, looking leader-like seems central to men's positive self-evaluations and might be important for men's CSEs because it communicates status and resources, which evolutionary theories suggest may hold value for men (e.g., Campbell & Wilbur, 2009; Li et al., 2002). Indeed, looking leader-like is strongly related to inferences of Power but also to inferences of Warmth.

Although people may learn to associate particular appearances with leadership and status through socialization, the predisposition to use scant facial information to evaluate others may form early in childhood (Cogsdill et al., 2014). Children's leadership ratings of faces likewise predicted politicians' electoral success in one study (Antonakis & Dalgas, 2009). These findings suggest that the perception of leadership (and possibly the selection of leaders) might arise from basic processes that operate without socialization. Furthermore, if leadership appearance relates to individuals' CSEs, then it might suggest that faces reflect internal traits associated with effective leadership behaviors (Hu et al., 2012). Thus, if children's face-based leadership judgments also relate to targets' CSEs and leadership success, then it would suggest that the processes involve elements that do not wholly rely on socialization, supporting their basic and adaptive nature. We therefore addressed this possibility in Study 2.

Study 2

In Study 2, we investigated whether children's leadership ratings would predict targets' CSEs and leadership success. If children's leadership ratings positively relate to targets' occupational status, leadership role, and CSEs like adults' ratings did in Study 1, then one may infer that leadership perception relies on a basic process that does not necessarily

require workplace socialization, and that both adults and children may use similar heuristics to perceive leadership from faces.

Method

      Participants. We recruited 878 children to participate in Study 2: 469 at University of Bonn during the German Unity Day celebrations, and another 409 from grades 3-6 at seven schools in a leading economic region of Germany. Of these, 12 children could not complete the task due to technical problems; we excluded another 47 from analysis for providing identical ratings in more than half of the task blocks. The excluded children ($M = 9.19$ years, $SD = 1.84$) were younger than those remaining in the sample [$M = 10.40$ years, $SD = 1.88$; $t(864) = 4.29$, $p < .001$, $r_{Effect\ Size} = .14$], likely because younger children may have had more trouble paying attention throughout the study. The two groups did not differ in gender [28 boys and 19 girls excluded, 460 boys and 359 girls in the final sample: χ^2 (1, $N = 819$) = 0.21, $p = .66$, $\phi = .02$] or whether they enjoyed the study [excluded children: $M = 4.45$, $SD = 0.77$; vs. remaining children: $M = 4.49$, $SD = 0.71$, on a scale from 1 (*Certainly not!*) to 5 (*Yes, sure!*); $t(863) = -0.45$, $p = .65$, $r_{Effect\ Size} = .02$].[6] All parents provided oral or written informed consent for their children's participation.

      Procedure. Children participated in groups of up to seven in the laboratory. An experimenter guided the children to sit in front of a laptop computer running Inquisit 3.0.5.0 (2011) and listen to the instructions through headphones while they simultaneously read them on the computer's screen. Before asking the children to rate the photos, we presented them with an animated story in which a ship's captain displayed several leadership behaviors and accomplished leadership tasks that included personnel selection, motivating the sailors on the ship, giving directions, and making decisions. These tasks were based on leadership functions

---

[6] One child did not complete this measure.

described by Mintzberg (1975) and the story was pre-tested and modified several times to assure that it matched the general level of understanding for children in the age range tested.

To keep the study brief, every child provided ratings for only six photos randomly chosen from either the 130 neutral or 130 impression-management photos used in Study 1 (see Table 1 for the average number, *k,* raters per judgment). The rating procedure consisted of seven blocks; in each block, children rated the same six faces presented in a different random order. The instructional story ended with the old captain retiring to a beautiful island and leaving his leadership position vacant. The animation consisted of drawings showing only the captain's back (thus, they did not provide any facial information that might have otherwise affected the children's subsequent ratings). The children were then asked in the first block to "*... recommend men to the old captain. The old captain will choose one of them as the new captain. Would you recommend this man?"* by rating each face on a 5-point scale ranging from 1 (*Certainly not!*) to 5 (*Yes, sure!*) with the anchors accompanied by cartoon faces displaying negative (frowning) and positive (smiling) faces, respectively. Although this task always occupied the first block, blocks two through six were randomly ordered and consisted of ratings aligned with the animated leadership story: *Does he always choose a good passage for the boat?*, *Is he good at explaining to the crew what to do?*, *Is he good at calming the crew down when they are afraid on dangerous trips?*, *Does he look like a captain?*, and *Does he look good?* (a measure of attractiveness), all using the same 5-point scale as in the first block. Finally, the children indicated whether they had recognized any of the faces and, if so, rated every face with regard to whether it was familiar to them in a final block; we removed data from recognized faces prior to analysis (2.5% of all trials). As in Study 1, we removed strings of identical ratings for a complete block and excluded *in toto* data from children who provided strings of identical ratings in more than half of the blocks (10.3% of all trials).

Estimation strategy. We related these judgments to the CSE, occupational status, and leadership role data for the targets while accounting for the Professional Coach dummy variable, as in Study 1. To guard against endogeneity bias, we followed the same IV-estimation procedures described in Study 1, except that we did not have to account for nesting by photoset because the photos were chosen randomly in Study 2.

Results

We found an acceptable degree of interrater agreement (all ICC2s > .70; see Table 1) and, thus, averaged each target's ratings across all of the children. As in Study 1, we excluded the attractiveness (here, *looks good*) rating from our analyses because it strongly correlated with the leadership ratings due to common source and method effects (all rs > .49, all ps < .001) (Antonakis et al., 2010; Podsakoff et al., 2012). The other five ratings concerned perceptions of leadership (*preselection as captain, looks like a captain, chooses a good passage, explains what to do, calms sailors down*); thus, we examined the appropriateness of combining them into one Leadership score by performing a principal components analysis with varimax rotation (e.g., Rule & Ambady, 2008). Indeed, the analysis returned a one-factor solution in both the neutral-photo and impression-management conditions, so we averaged the individual ratings into a single Leadership composite (see Table 5). We also calculated ICC scores for the Leadership composite after averaging the five ratings within each rater; the resulting ICC2 scores revealed acceptable values for both photo types (see Table 1). Because the ratings of the targets correlated strongly across the impression-management and neutral photos (all rs > .59, all ps < .001), we averaged them into a single index for our main analyses, as above (see Table 6 for the means, standard deviations, and correlations for all variables).

The associations between Leadership and the targets' leadership roles ($B = 0.81$, $SE = 0.36$, $p = .02$; Table 4, Model 3a) and occupational status ($B = 1.54$, $SE = 0.75$, $p = .04$; Model 5a) were positive and significant, reinforcing our support for Hypothesis 1. The association

between Leadership and leadership role remained significant when controlling for

Professional Coach status ($B = 0.78$, $SE = 0.36$, $p = .03$; Model 3b) but the relation with

occupational status became marginally significant ($B = 1.43$, $SE = 0.73$, $p = .051$; Model 5b).

Leadership also positively associated with CSEs ($B = 0.34$, $SE = 0.11$, $p = .001$; Model 1a),

regardless of whether Professional Coach status was included ($B = 0.34$, $SE = 0.11$, $p = .002$;

Model 1b), consistent with Hypothesis 2.

Instrumental-variable mediation analyses. CSEs significantly predicted leadership

role ($B = 2.99$, $SE = 1.23$, $p = .02$) and occupational status ($B = 5.82$, $SE = 2.37$, $p = .01$),

regardless of Professional Coach status (leadership role: $B = 2.92$, $SE = 1.24$, $p = .02$;

occupational status: $B = 5.50$, $SE = 2.35$, $p = .02$); see Table 4, Models 4a, 6a, 4b, and 6b,

respectively). Based on these results, we conducted several mediation analyses using a bias-

corrected bootstrapping procedure with 10,000 bootstraps to derive the CIs for the indirect

effect.

Supporting Hypothesis 3, we found that CSEs again mediated the positive association

between Leadership and both success criteria (leadership role: indirect effect = 0.91, $SE =$

0.31, 95% CI [0.30, 1.49]; occupational status: indirect effect = 1.76, $SE = 0.65$, 95% CI

[0.48, 3.16]; see Table 4, Models 2a, 4a, 2a, and 6a, respectively), regardless of whether we

accounted for Professional Coach status (all indirect effects ≥ 0.87, all SEs ≤ 0.63, all 95%

CIs [0.24, 3.01]). Moreover, the correlations between the disturbances of CSEs with

leadership role ($\Psi = -0.39$, $SE = 0.21$, $p = .059$) and occupational status ($\Psi = -0.65$, $SE =$

0.40, $p = .102$) approached significance, indicating that the CSEs may be endogenous. We

therefore conducted a stronger test for endogeneity and compared the likelihood ratio of the

instrumental variable model to one in which we constrained the disturbances of CSE-

leadership role and CSEs-occupational status to zero (Antonakis et al., 2010). The likelihood

ratio test was significant, $\chi^2(2) = 6.40$, $p = .04$ (regardless of whether Professional Coach

status was included, $\chi^2(2) = 6.02$, $p = .049$), again supporting the endogeneity of the CSEs and

the need for IV-estimation to obtain consistent estimates (Antonakis et al., 2010).

Overidentification tests indicated that the mediation models fit the data well [$\chi^2(2) = 1.17$, $p =$

.56; with Professional Coach dummy: $\chi^2(2) = 1.12$; $p = .57$]. In addition, none of the

modification indices exceeded the minimum value, suggesting that our models fit the data

well.

      Exploratory analyses with adults' trait inferences. Similar to Study 1, the

children's Leadership judgments strongly related to the adults' trait inferences of Warmth

[$r(128) = .67$, $p < .001$] and moderately to their judgments of Power [$r(128) = .45$, $p < .001$].

Discussion

      The results of Study 2 showed again that men who looked more like leaders had

achieved higher positions of leadership and greater occupational status. More important, the

men's self-reported CSEs mediated this association. Further, children's perceptions of

leadership strongly related to adults' inferences of Warmth, and moderately related to adults'

inferences of Power. The similar pattern of results across children and adults in Studies 1 and

2 suggests that both groups use similar heuristics when perceiving leadership from faces,

supporting the possibility that such judgments may rely on basic processes that operate

without work experience or workplace socialization. In addition, the findings suggest that

inferences based on scant facial information can predict important real-world phenomena like

increased occupational status, leadership role attainment, and CSEs. Indeed, the association

between CSEs and success documented in previous work (e.g., Judge et al., 2003; Judge,

Hurst et al., 2009) supports the possibility that perceptions of leadership from the face might

reflect the targets' actual leadership ability.

General Discussion

      Here, we found that perceptions of leadership based on the face correlated with men's

success at work. Both children's and adults' judgments of leadership from men's faces

significantly corresponded to the leadership roles that they occupied in their jobs and to their

occupational status. Critically, the targets' CSEs mediated these associations. Their apparent

leadership ability therefore related to their internal traits (i.e., CSEs), which in turn correlated

with their success in leadership roles within their professional organizations. This suggests

that superficial perceptions of leadership may index abilities that reliably and meaningfully

predict legitimate leadership behaviors. Moreover, our observation that judgments made by

both children and adults lead to similar conclusions suggests that the processes by which

leadership is expressed and perceived might have very basic roots that do not require

workplace socialization.

Adults' and children's perceptions of leadership revealed small-to-medium positive

associations with occupational status and leadership role. These findings align with the results

of previous research showing that first impressions of faces predict leadership outcomes in

business, politics, law, and the military (e.g., Mueller & Mazur, 1996; Rule & Ambady, 2008,

2011; Todorov et al., 2005).

The results of both Studies 1 and 2 underscored the importance of perceptions of

leadership from the face for men's self-concepts. Evolutionary models of mating posit that

status and resources promote competitive advantages among men, and evolutionary models of

self-concept have shown that status-related traits are integral for men (e.g., Buss, 1989;

Campbell & Wilbur, 2009; Li et al., 2002). Consistent with these ideas, having a face that

signals status and resources (as through conveying impressions of leadership, for example)

may encourage positive self-concepts in men.

Moreover, the associations between perceptions of leadership, CSEs, and success

appeared to be quite robust. We also found remarkable consistency between perceptions made

by both the children in Study 2 and the adults in Study 1. In line with findings showing that

adults' and children's perceptions of faces follow similar processes with relatively equivalent

outcomes (Antonakis & Dalgas, 2009; Cogsdill et al., 2014), the association between success

and facial appearance seems to generalize for both children and adult observers. This

consistency points to the possibility that people may perceive leadership quickly and without relevant experience due to evolved adaptive mechanisms by which the face functions as a diagnostic tool (Antonakis & Dalgas, 2009; Ballew & Todorov, 2007; Spisak, Dekker et al., 2012; Spisak, Homan et al., 2012). Specifically, the similar responses of adults and children suggest that workplace experiences may not be necessary to learn leadership prototypes. Rather, leadership prototypes could stem from more basic processes potentially honed over the course of human evolution (e.g., Rule, Moran, Freeman, Whitfield-Gabrieli, Gabrieli, & Ambady, 2011; Spisak, Homan et al., 2012), via an innate face template (Reid, Dunn, Young, Amu, Donovan, & Reissland, 2017), or from early rapid learning (Slater & Quinn, 2001). For instance, human groups that could quickly select effective leaders might have enjoyed survival advantages compared to groups that could not (Van Vugt et al., 2008). Honestly advertising leadership ability through one's face might have thus facilitated leader emergence and enabled group coordination. We offer these speculations tentatively, because they involve processes that would be difficult to test directly, requiring substantial future work before permitting strong conclusions. Moreover, self-fulfilling prophecies could explain why appearance, CSEs, and success relate, given that individuals who look like leaders may be treated like leaders and the provided leadership experiences allow them to develop leadership ability, further reinforcing the initial impression (Antonakis & Eubanks, 2017; Rule & Ambady, 2011; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Thus, an individual might not possess leadership ability at the outset but cultivate it as a byproduct of fitting an apparent leader prototype.

Likewise, Judge, Piccolo, and Kosalka (2009) theorized that individual differences may promote the diverse skills needed for leadership, thereby enhancing the survival of a group. Although the advantages of being a leader (e.g., to have more resources, such as opportunities for procreation) and having high CSEs may seem obvious, one might question why anyone would choose to be a follower? First, followers may realize that their traits do not

favor their ascension into a leadership position; thus, they increase their chances of survival by following others (Judge et al., 2009). Second, leaders often must delegate tasks to followers (e.g., to supplement their own expertise), which can include similar rewards. Following, then, does not purely mean subjugation; rather, it can meaningfully facilitate an individual's survival and prosperity by cooperating with others who have different skills.

Strengths and Limitations

Despite many important studies documenting the association between first impressions of leadership and various real-life outcomes, this is the first empirical investigation of how these associations may manifest within the workplace. Moreover, most of the prior research on this topic has examined individual elite leaders across organizations (see Re & Rule, 2017; Rule & Tskhay, 2014). In contrast, the present work investigated how perceptions of leadership within the rank-and-file members of an organization also predict their success. Thus, both the targets and participants were real employed adults, rather than undergraduates lacking work experience, and we observed parallel results when sampling children. This allows us to generalize the findings beyond a specific company or profession, demonstrating high ecological and external validity. Moreover, using multiple criteria (occupational status and leadership role occupancy) and conducting a multi-sample study of raters further increases the generalizability of our findings. In addition, we found convergent results across multiple standardized stimuli of targets (i.e., photos with a neutral facial expression and an impression-management facial expression), allowing us to control for different gestures and postures in pictures. This not only speaks well to the internal validity of our research, but also helps to efficiently demonstrate the face's role in these judgments (rather than its ephemeral social presentation; see Rule, Krendl, Ivcevic, & Ambady, 2013; Todorov & Porter, 2014).

This research also has several limitations, however. For instance, our cross-sectional design challenges our capacity to draw causal inferences, even though the proposed theory supports our model (Antonakis, 2011; Judge, Hurst et al., 2009; Spisak, Dekker et al., 2012).

Moreover, longitudinal studies found that CSEs predicted individuals' work success and growth trajectories (Judge & Hurst, 2008; Judge, Hurst et al., 2009). Thus, the hypothesized mediation seems sensible, despite the cross-sectional design, because we can assume that the variables we studied exist in a sort of stable equilibrium: facial appearance predicts CSEs, which predict success at work (Fischer, Dietz & Antonakis, 2017).

That said, both facial appearance and CSEs may be somewhat plastic. Although many regard CSEs as stable, they show malleability just as other traits do (e.g., neuroticism, self-esteem; Nübold et al., 2013; Orth et al., 2010; Roberts, Walton, & Viechtbauer, 2006; Wu & Griffin, 2012). Similarly, experiencing success at work could positively influence one's CSEs and traits like CSEs could influence appearance (e.g., Dorian Gray effects; see Zebrowitz, 1997). Although previous research showing that facial cues predict leadership outcomes decades later casts doubt on the latter (Mueller & Mazur, 1996, Rule & Ambady, 2011; see also Zebrowitz, Olson, & Hoffman, 1993), we recognize that CSEs and success may influence each other bidirectionally (see also Roberts, Caspi, & Moffitt, 2003). Thus, we conducted mediation analyses using an instrumental variable estimator to guard against endogeneity bias and a limited-information maximum likelihood estimator, which is partially robust to weak instruments (e.g., age in our studies) and therefore more reliable in these circumstances than a two-step least squares estimator (e.g., Staiger & Stock, 1997; Stock, Wright, & Yogo, 2002).

Yet, the variables we investigated (facial appearance, CSEs, and success) might still all arise from omitted variables or a common cause that we did not measure (e.g., general fitness; Antonakis & Eubanks, 2017). For instance, indicators of general fitness such as facial symmetry, intelligence, and extraversion all correlate (Antonakis & Eubanks, 2017). In addition, omitted variables such as intelligence and motivation both relate to CSEs and success, therefore potentially accounting for the associations we found (e.g., intelligence and motivation can antecede and succeed CSEs, respectively; Ferris, Rosen, Johnson, Brown, Risavy, & Heller, 2011; Judge, Hurst et al., 2009). Moreover, we cannot exclude the

possibility that children's and adults' converging leadership perceptions both result from early socialization, particularly as the children were already roughly 10 years old (similar to the ages of participants in Antonakis & Dalgas's, 2009, research). Experiences with one's family, school, or media (e.g., TV, movies) might foster leadership prototypes in children, though previous research found that the predictions of political leaders' electoral success did not depend on age (regardless of whether a 5-year-old child or a 70-year-old adult made the judgment; Antonakis & Dalgas, 2009). Because children as young as five years do not have much experience with workplace leadership, our concordant results for children and adults appear even more remarkable, and align with research showing that preschool children reliably attributed trustworthiness and competence to faces (Cogsdill et al., 2014).

In addition, although we excluded all ratings of participants who indicated that they recognized a target face, the handball player participants in Study 1 might have given unconsciously biased judgments resulting from mere exposure to the target faces (though replication of those results with the child sample in Study 2 discourages this possibility).

Finally, because the CSES measures CSEs generally, we could not investigate whether these processes might have varied across its four subcomponents (i.e., emotional stability, self-esteem, generalized self-efficacy, and locus of control). Measuring CSEs directly, rather than as a multidimensional construct, can be problematic (Chang et al., 2012; Johnson, Rosen, & Levy, 2008). For instance, multiple CSE components could differentially affect work outcomes, as they may have different antecedents (Johnson, Rosen, & Levy, 2008). Future research may therefore benefit from examining the relation between appearance and these specific components. Modeling them in future research would also help to determine the robustness of our results.

Directions for Future Research

Certainly, future studies could extend the present findings in a number of ways. For instance, because of the low frequency of female coaches in A-level workshops of the

German Handball Association, we only investigated male targets. This leaves open the question of how gender might influence the internal traits that account for the association between perceptions of leadership and leadership outcomes observed in other work (e.g., Chiao, Bowman, & Gill, 2008; Rule & Ambady, 2009). Indeed, previous studies might suggest that the association between perceptions of leadership, CSEs, and professional success may be deeper and more complex for female leaders (e.g., Hogue & Lord, 2007; Joshi, Son, & Roh, 2015; Silberzahn & Menges, in press).

Similarly, given past research showing that leadership perceptions can vary based on cultural values (e.g., Rule et al., 2010; see also Abdalla & Al-Homoud, 2001; House, Javidan, Hanges, & Dorfman, 2002), extending these tests to targets and participants from diverse cultural backgrounds might help to advance the present findings as well. Consistent findings across such diverse samples might especially help to address the currently speculative possibility that evolutionary processes might partially explain the phenomena we observed.

Future studies might also use a longitudinal design to investigate whether leadership- or status-related impressions might positively influence the development of one's self-concept. These might not only include CSEs, but also other self-concept aspects, such as one's self-perception as a leader and identification with leadership roles.

Finally, although we found that perceptions of leadership related to targets' internal CSEs and demonstrated leadership success, one's appearance does not always convey elements of truth (see Jussim, Crawford, & Rubinstein, 2015, for review). Rather, facial appearance might unproductively bias leader selection and future research could help to discern when facial appearance facilitates or misleads individuals in their decisions to follow a particular leader.

Practical Implications

These findings may have practical implications for organizational assessment systems presently used to make decisions about employment and promotion. Decision makers

generally seek to minimize the influence of job-irrelevant human capital (Baruch &

Bozionelos, 2011). Yet, ample evidence shows that employers favor people with particular

appearances no matter how much job-relevant information they have available to them

(Hosoda et al., 2003; Rule, Bjornsdottir, Tskhay, & Ambady, 2016). Thus, decision makers

should ensure that organizational assessment systems emphasize the knowledge, skills, and

abilities needed for a position. Moreover, interviewers should focus on the relevant behavior

of a candidate, which helps to protect against erroneously hiring (or dismissing) someone

because he or she simply "looks right" for a particular position.

Conclusion

The present research underscores the importance of facial appearance for career

success and leadership role occupancy. Adults' and children's judgments of faces correlated

with targets' occupational status and leadership role at their place of work. CSEs mediated

this association, suggesting that having a face that conveys status may contribute to a positive

self-concept in men. Consistency between the perceptions of adults and children suggests that

the mechanism underlying these relations do not rely on socialization within an organizational

framework. Thus, what shows on the "outside" of a person may, in some instances, match a

bit of what he or she holds on the "inside," potentially forecasting the efficacy of that

individual as a leader.

References

Abdalla, I. A., & Al-Homoud, M. A. (2001). Exploring the implicit leadership theory in the Arabian Gulf states. *Applied Psychology*, *50*, 506-531.

Antonakis J. (2011). Predictors of leadership: The usual suspects and the suspect traits. In Bryman A., Collinson D., Grint K., Jackson B. & Uhl-Bien M. (Eds.), *Sage Handbook of Leadership* (pp. 269-285). Thousand Oaks: Sage Publications. http://dx.doi.org/10.1126/science.1167748

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, (in press). http://dx.doi.org/10.1016/j.leaqua.2017.01.006

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*, 1086-1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play!. *Science, 323*, 1183.

Antonakis, J., & Eubanks, D. L. (2017). Looking leadership in the face. *Current Directions in Psychological Science*, https://doi.org/10.1177/0963721417705888.

Ballew, C. C., II, & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences, 104*, 17948–17953. http://dx.doi.org/10.1073/pnas.0705435104

Bamberger, P. A., & Pratt, M. G. (2010). Moving forward by looking back: Reclaiming unconventional research contexts and samples in organizational scholarship. *Academy of Management Journal, 53*, 665-671. http://dx.doi.org/10.5465/AMJ.2010.52814357

Baruch, Y. & Bozionelos, N. (2011). Career issues. In S. Zedeck (Ed.), *APA Handbook of Industrial and Organizational psychology: Selecting and Developing Members of the Organization* (Vol. 2, pp. 67-113). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/12170-003

Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E.

(2016). Statistical control in correlational studies: 10 essential recommendations for

organizational researchers. *Journal of Organizational Behavior, 37*, 157 – 167.

http://dx.doi.org/10.1002/job.2053

Berry, D. S. (1990). Taking people at face value: Evidence for the kernel of truth hypothesis.

Social Cognition, 8, 343–361. http://dx.doi.org/10.1521/soco.1990.8.4.343

Betzig, L. (1993). Sex, succession, and stratification in the first six civilizations: How

powerful men reproduced, passed power on to their sons, and used power to defend

their wealth, women, and children. In L. Ellis (Ed.), *Social stratification and

socioeconomic inequality* (Vol. 1, pp. 37–74). Westport, CT: Praeger.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal

of Personality and Social Psychology, 66,* 5-20. http://dx.doi.org/10.1037/0022-

3514.66.1.5

Blickle, G., Fröhlich, J., Ehlert, S., Pirner, K., Dietl, E., Hanes T. J. & Ferris, G. R. (2011).

Socioanalytic theory and work behavior: Roles of work values and political skill in job

performance and promotability assessment. *Journal of Vocational Behavior, 78*, 136-

148. http://dx.doi.org/10.1016/j.jvb.2010.05.010

Blickle, G., Witzki, A., & Schneider, P.B. (2009). Mentoring support and power: A three year

predictive field study on protégé networking and career success. *Journal of Vocational

Behavior, 74*, 181-189. http://dx.doi.org/10.1016/j.jvb.2008.12.008

Bliese, P. D. (1998). Group size, ICC values, and group-level correlations: A simulation.

Organizational Research Methods, 1, 355-373.

https://doi.org/10.1177/109442819814001

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications

for data aggregation and analyses. In K. J. Klein & S. W. J. Kozlowski (Eds.),

Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions (pp. 349– 381). San Francisco: Jossey-Bass.

Bøggild, T., & Laustsen, L. (2016). An intra-group perspective on leader preferences: Different risks of exploitation shape preferences for leader facial dominance. *Leadership Quarterly, 27*, 820–837. http://dx.doi.org/10.1016/j.leaqua.2016.09.003

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences, 12*, 1-49. http://dx.doi.org/10.1017/S0140525X00023992

Buss, D. M. (2005). *Handbook of evolutionary psychology*. Hoboken, NJ: Wiley.

Campbell, L., & Wilbur, C.J. (2009). Are the traits we prefer in potential mates the traits they value in themselves? An analysis of sex differences in the self-concept. *Self and Identity, 8*, 418-446. http://dx.doi.org/10.1080/15298860802505434

Chang, C.-H., Ferris, D. L., Johnson, R. E., Rosen, C. C., & Tan, J. A. (2012). Core self-evaluations: A review and evaluation of the literature. *Journal of Management, 38*, 81-128. http://dx.doi.org/10.1177/0149206311419661

Chiao, J. Y., Bowman, N. E., & Gill, H. (2008). The political gender gap: Gender bias in facial inferences that predict voting behavior. *PLoS ONE, 3*, 3666. http://dx.doi.org/10.1371/journal.pone.0003666

Cogsdill, E., Todorov, A., Spelke, E., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological Science, 25*, 1132-1139. http://dx.doi.org/10.1177/0956797614523297

Dietl, E., Meurs, J. A., & Blickle, G. (2017). Do they know how hard I work? Investigating how implicit/explicit achievement orientation, reputation, and political skill affect occupational status. *European Journal of Work and Organizational Psychology, 26*, 120-136. http://dx.doi.org/10.1080/1359432X.2016.1225040

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*, 285–290. http://dx.doi.org/10.1037/h0033731

Ehrenspiel, F., & Strahler, K. (2012). *Psychoneuroendocrinology of Sport and Exercise*. London: Routledge.

Elgar, M. A. (2016). Leader selection and leadership outcomes: Height and age in a sporting model. *The Leadership Quarterly*, *27*, 588-601. https://doi.org/10.1016/j.leaqua.2015.12.005

Ferris, D. L., Rosen, C. R., Johnson, R. E., Brown, D. J., Risavy, S. D., & Heller, D. (2011). Approach or avoidance (or both?): Integrating core self-evaluations within an approach/avoidance framework. *Personnel Psychology*, *64*, 137-161. https://doi.org/10.1111/j.1744-6570.2010.01204.x

Fischer, T., Dietz, J., & Antonakis, J. (2017). Leadership process model: A review and synthesis. *Journal of Management, 43*, 1726-1753. http://dx.doi.org/10.1177/0149206316682830

Friedman, H., & Zebrowitz, L. A. (1992). The contribution of typical sex differences in facial maturity to sex role stereotypes. *Personality and Social Psychology Bulletin, 18*, 430–438.

Ganzeboom, H. B. G. (2010, May). *A new international socio-economic index (ISEI) of occupational status for the international standard classification of occupation 2008 (ISCO-08) constructed with data from the ISSP 2002-2007*. Paper presented at the Annual Conference of International Social Survey Programme, Lisbon.

Ganzeboom, H. B. G., & Treiman, D. J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in cross-national comparison. A European working book for demographic and socio-economic variab*les. (pp. 159-193). New York: Kluwer Academic/Plenum Publishers.

Harms, P., Han, G., & Chen, H. (2012). Recognizing leadership at a distance: A study of

    leader effectiveness across cultures. *Journal of Leadership and Organizational*

    Studies, 19, 164-172. http://dx.doi.org/10.1177/1548051812436812

Harter, S. (2006). The development of self-esteem. In M. Kernis (Ed.), *Self-esteem issues and*

    answers: A sourcebook of current perspectives (pp. 144–150). New York: Psychology

    Press.

Hartmann, M. (2002). *Der Mythos der Leistungseliten [The myth of the achievement elites.*

    Top careers and social background in economy, politics, justice administration, and

    science]. Frankfurt: Campus.

Hiller, N. J., & Hambrick, D. C. (2005). Conceptualizing executive hubris: The role of

    (hyper-) core self-evaluations in strategic decision-making. *Strategic Management*

    Journal, 26, 297−319.

Hogue, M., & Lord, R. G. (2007). A multilevel, complexity theory approach to understanding

    genderbias in leadehip. *The Leadership Quarterly, 18*, 370–390.

    doi:10.1017/S0140525X0999152X

Hosoda, M., Stone-Romero, E. F., & Coats, G. (2003). The effects of physical attractiveness

    on job-related outcomes: A meta-analysis of experimental studies. *Personnel*

    Psychology, 56, 431-462. http://dx.doi.org/10.1111/j.1744-6570.2003.tb00157.x

House, R., Javidan, M., Hanges, P., & Dorfman, P. (2002). Understanding cultures and

    implicit leadership theories across the globe: an introduction to project GLOBE.

    Journal of World Business, *37*, 3-10.

Hu, J., Wang, Z., Liden, R. C. & Sun, J. (2012). The influence of leader core self-evaluation

    on follower reports of transformational leadership. *The Leadership Quarterly, 23*, 860-

    868. http://dx.doi.org/10.1016/j.leaqua.2012.05.004

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job

    performance. *Psychological Bulletin, 96*, 72–98. http://dx.doi.org/10.1037/0033-

2909.96.1.72

Inquisit 3.0.5.0 [Computer software]. (2011). Seattle, WA: Millisecond Software.

Johnson, R. E., Rosen, C. C., & Levy, P. E. (2008). Getting to the core of core self-
evaluation: a review and recommendations. *Journal of Organizational Behavior*, *29*,
391-413. http://dx.doi.org/10.1002/job.514

Joshi, A., Son, J., & Roh, H. (2015). When can women close the gap? A meta-analytic test of
sex differences in performance and rewards, *Academy of Management Journal, 58*,
1516-1545. http://dx.doi.org/10.5465/amj.2013.0721

Judge, T. A., Bono, J. E., Erez, A., & Locke, E. A. (2005). Core self-evaluations and job and
life satisfaction: The role of self-concordance and goal attainment. *Journal of Applied
Psychology, 90*, 257-268. http://dx.doi.org/10.1037/0021-9010.90.2.257

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A
qualitative and quantitative review. *Journal of Applied Psychology, 87*, 765-780.
http://dx.doi.org/10.1037/0021-9010.87.4.765

Judge, T. A., Cable, D. M., Boudreau, J. W., & Bretz, R. D. (1995). An empirical
investigation of the predictors of executive career success. *Personnel Psychology, 48*,
485-519. http://dx.doi.org/10.1111/j.1744-6570.1995.tb01767.x

Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative
review and test of theoretical propositions. *Journal of Applied Psychology, 89*, 542-
552. http://dx.doi.org/10.1037/0021-9010.89.3.542

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale
(CSES): Development of a measure. *Personnel Psychology, 56*, 303-331.
http://dx.doi.org/10.1111/j.1744-6570.2003.tb00152.x

Judge, T. A., & Hurst, C. (2008). How the rich (and happy) get richer (and happier):
Relationship of core self-evaluations to trajectories in attaining work success. *Journal
of Applied Psychology, 93*, 849–863. http://dx.doi.org/10.1037/0021-9010.93.4.849

Judge, T. A., Hurst, C., & Simon, L. N. (2009). Does it pay to be smart, attractive, or confident (or all three)?: Relationships among general mental ability, physical attractiveness, core self-evaluations, and income. *Journal of Applied Psychology, 94*, 742-755. http://dx.doi.org/10.1037/a0015497

Judge, T. A., Piccolo, R. F., & Kosalka, T. (2009). The bright and dark sides of leader traits: A review and theoretical extension of the leader trait paradigm. *The Leadership Quarterly, 20*, 855–875.

Jussim, L., Crawford, J. T., & Rubinstein, R. S. (2015). Stereotype (in)accuracy in perceptions of groups and individuals. *Current Directions in Psychological Science, 24*, 490–497.

Krendl, A. C., Rule, N. O., & Ambady, N. (2014). Does aging impair first impression accuracy? Differentiating emotion recognition from complex social inferences. *Psychology and Aging*, *29*, 482-490. http://dx.doi.org/10.1037/a0037146

Langlois J.H., Kalakanis L., Rubenstein A.J., Larson A., Hallam M., & Smoot M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*, 390-423. http://dx.doi.org/10.1037/0033-2909.126.3.390

Laustsen, L., & Petersen, M. B. (2017). Perceived conflict and leader dominance: Individual and contextual factors behind preferences for dominant leaders. *Political Psychology*. http://doi.org/10.1111/pops.12403

Lawson, C., Lenz, G. S., Baker, A., & Myers, M. (2010). Looking like a winner: Candidate appearance and electoral success in new democracies. *World Politics, 62*, 561-593. http://dx.doi.org/10.1017/S0043887110000195

Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. W. (2002). The necessities and luxuries of mate preferences: testing the tradeoffs. *Journal of Personality and Social Psychology, 82*, 947-955. http://dx.doi.org/10.1037/0022-3514.82.6.947

Li, W.-D., Arvey, R. D. & Song, Z. (2011). The influence of general mental ability, self-

esteem and family socioeconomic status on leadership role occupancy and leader advancement: The moderating role of gender. *The Leadership Quarterly, 22*, 520-534. http://dx.doi.org/10.1016/j.leaqua.2011.04.009

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. http://dx.doi.org/10.1016/j.evolhumbehav.2006.09.002

Lord, R. G., Foti, R. J. & De Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance, 34*, 343-378. http://dx.doi.org/10.1016/0030-5073(84)90043-6

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*, 30-46. http://dx.doi.org/10.1037/1082-989X.1.1.30

Martin, D. S. (1978). Person perception and real-life electoral behavior. *Australian Journal of Psychology, 30*, 255. http://dx.doi.org/10.1080/00049537808256378

Miller, D. C., & Salkind, N. J. (2002). *Handbook of research design and social measurement*. London, UK: Sage.

Mintzberg, H. (1975). The manager's job: Folklore and fact. *Harvard Business Review, 53*, 49-61.

Mueller, U. & Mazur, A. (1996). Facial dominance of West Point cadets as predictor of later military rank. *Social Forces, 74*, 823-850. http://dx.doi.org/10.2307/2580383

Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nübold, A., Muck, P., & Maier, G.W. (2013). A new substitute for leadership? Followers' state core self-evaluations. *The Leadership Quarterly, 24*, 29-44 doi: 10.1016/j.leaqua.2012.07.002

Orth, U., Trzesniewski, K. H., & Robins, R. W. (2010). Self-esteem development from young

   adulthood to old age: A cohort-sequential longitudinal study. *Journal of Personality and*

   Social Psychology, 98, 645-658. http://dx.doi.org/10.1037/a0018769

Penton-Voak, I.S., Pound, N., Little, A.C., & Perrett, D.I. (2006). Personality judgements

   from natural and composite facial images: More evidence for a "kernel of truth" in

   social perception. *Social Cognition, 24*, 607–640.

   http://dx.doi.org/10.1521/soco.2006.24.5.607

Re, D. E., Hunter, D. W., Coetzee, V., Tiddeman, B. P., Xiao, D. X., 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,* e80957.

   http://dx.doi.org/10.1371/journal.pone.0080957

Re, D. E., & Rule, N. O. (2015). CEO facial appearance, firm performance, and financial

   success. In M. Fetscherin (Ed.) *CEO Branding: Meaning, Measuring, Managing* (pp.

   219-238). New York, NY: Routledge.

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

Reid, V. M., Dunn, K., Young, R.Y., Amu, J., Donovan, T., & Reissland, N. (2017). The

   human fetus preferentially engages with face-like visual stimuli. *Current Biology, 27*,

   1825–1828. http://dx.doi.org/10.1016/j.cub.2017.05.044

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configura-

   tions affect strategic choices in Trust Games with or without information about past be-

   havior. *PloS ONE*, *7*, e34293. https://doi.org/10.1371/journal.pone.0034293

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality

   development in young adulthood. *Journal of Personality and Social Psychology, 84*,

   582. http://dx.doi.org/10.1037/0022-3514.84.3.582

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in

    personality traits across the life course: A meta-analysis of longitudinal studies.

    Psychological Bulletin, 132, 1-25. http://dx.doi.org/10.1037/0033-2909.132.1.1

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart,

    & Winston.

Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from Chief Executive

    Officers' appearance predict company profits. *Psychological Science, 19*, 109-111.

    http://dx.doi.org/10.1111/j.1467-9280.2008.02054.x

Rule, N. O., & Ambady, N. (2009). She's got the look: Inferences from female chief

    executive officers' faces predict their success. *Sex Roles, 61,* 644-652.

    http://dx.doi.org/10.1007/s11199-009-9658-9

Rule, N. O., & Ambady, N. (2011). Judgments of power from college yearbook photos and

    later career success. *Social Psychological and Personality Science, 2*, 154-158.

    http://dx.doi.org/10.1177/1948550610385473

Rule, N. O., Ambady, N., Adams, R. B., Jr., Ozono, H., Nakashima, S., Yoshikawa, S., &

    Watabe, M. (2010). Polling the face: Prediction and consensus across cultures. *Journal*

    of Personality and Social Psychology, 98, 1-15. http://dx.doi.org/10.1037/a0017673

Rule, N. O., Bjornsdottir, R. T., Tskhay, K. O., & Ambady, N. (2016). Subtle perceptions of

    male sexual orientation influence occupational opportunities. *Journal of Applied*

    Psychology, 101, 1687-1704. doi:10.1037/apl0000148

Rule, N. O., Moran, J. M., Freeman, J. B., Whitfield-Gabrieli, S., Gabrieli, J. D. E., &

    Ambady, N. (2011). Face value: Amygdala response reflects the validity of first

    impressions. *NeuroImage, 54,* 734-741.

    http://dx.doi.org/10.1016/j.neuroimage.2010.07.007

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. http://dx.doi.org/10.1037/a0031050

Rule, N. O., & Tskhay, K. O. (2014). The influence of economic context on the relationship between Chief Executive Officer facial appearance and company profits. *The Leadership Quarterly, 25*, 846-854. http://dx.doi.org/10.1016/j.leaqua.2014.01.001

Samochowiec, J., Wänke, M., & Fiedler, K. (2010). Political ideology at face value. *Social Psychological & Personality Science, 1*, 206-213. http://dx.doi.org/10.1177/1948550610372145

Schooler, C., & Schoenbach, C. (1994). Social class, occupational status, occupational self-direction, and job income: A cross-national examination. *Sociological Forum, 9*, 431-458. http://dx.doi.org/10.1007/BF01466317

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C. & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences, 276*, 575-584. http://dx.doi.org/10.1098/rspb.2008.1177

Shondrick, S. J., & Lord, R. G. (2010). Implicit leadership and followership theories: Dynamic structures for leadership perceptions, memory, and leader–follower processes. In G. P. Hodgkinson & J.K. Ford (Eds.), *International Review of Industrial and Organizational Psychology, 25*. (pp. 1−33). http://dx.doi.org/10.1002/9780470661628.ch1

Short, L.A., Mondloch, C.J., McCormick, C.M., Carré, J.C., Ma, R., Fu, G., & Lee, K. (2012). Detection of propensity for aggression based on facial structure irrespective of face race. *Evolution and Human Behavior, 33*, 121-129. http://dx.doi.org/10.1016/j.evolhumbehav.2011.07.002

Silberzahn & Menges, (in press). Reading the face of a leader: woman with low facial

    masculinity are perceived as competitive. *Academy of Management Discoveries*.

Slater, A., & Quinn, P. C. (2001). Face recognition in the newborn infant. *Infant and Child

    Development, 10*, 21-24. http://dx.doi.org/10.1002/icd.241

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal

    behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and

    Social Psychology, 35*, 656–666.

Spisak, B. R., Dekker, P. H., Kruger, M., & Van Vugt, M. (2012). Warriors and

    peacekeepers: Testing a biosocial implicit leadership hypothesis of intergroup

    relations using masculine and feminine faces. *PloS ONE,* 7, e30399.

    http://dx.doi.org/10.1371/journal.pone.0030399

Spisak, B. R., Homan, A. C., Grabo, A. & Van Vugt, M. (2012). Facing the situation: Testing

    a biosocial contingency model of leadership in intergroup relations using masculine

    and feminine faces. *The Leadership Quarterly, 23*, 273-280.

    http://dx.doi.org/10.1016/j.leaqua.2011.08.006

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments.

    Econometrica, 65, 557-586. http://dx.doi.org/10.2307/2171753Stock, J. H., Wright, J.

    H., & Yogo, M. (2002). A survey of weak instruments and weak identification in

    generalized method of moments. *Journal of Business & Economic Statistics, 20*, 518-

    529. https://doi.org/10.1198/073500102288618658

Stumpp, T., Muck, P., Hülsheger, U., Judge, T. A. & Maier, G. (2010). Core self-evaluations

    in Germany: Validation of a German measure and its relationships with career success.

    Applied Psychology: An International Review, 59, 674-700.

    http://dx.doi.org/10.1111/j.1464-0597.2010.00422.x

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence

from faces predict election outcomes. *Science, 308*, 1623-1626.

http://dx.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*, 519-545. https://doi.org/10.1146/annurev-psych-113011-

143831

Todorov, A., & Porter, J. (2014). Misleading first impressions: Different for different facial

images of the same person. *Psychological Science, 25*, 1404-1417.

http://dx.doi.org/10.1177/0956797614532474

Treiman, D. J. (1977). *Occupational prestige in comparative perspective*. New York:

Academic Press.

Trichas, S., & Schyns, B. (2012). The face of leadership: Perceiving leaders from facial

expressions. *The Leadership Quarterly, 23*, 545–566.

http://dx.doi.org/10.1016/j.leaqua.2011.12.007

Tskhay, K. O., Xu, H., & Rule, N. O. (2014). Perceptions of leadership success from

nonverbal cues communicated by orchestra conductors. *The Leadership Quarterly, 25,*

901-911. *http://dx.doi.org/10.1016/j.leaqua.2014.07.001*

Van Vugt, M., Hogan, R., Kaiser, R. (2008). Leadership, followership, and evolution: Some

lessons from the past. *American Psychologist, 63*, 182-196.

http://dx.doi.org/10.1037/0003-066X.63.3.182

Wilson, D. S., Van Vugt, M., & O'Gorman, R. (2008). Multilevel selection theory and major

evolutionary transitions: Implications for psychological science. *Current Directions in

Psychological Science, 17*, 6-9. http://dx.doi.org/10.1111/j.1467-8721.2008.00538.x

Wu, C-H., & Griffin M. A. (2012). Longitudinal Relationships between Core Self-

Evaluations and Job Satisfaction. *Journal of Applied Psychology, 97*, 331-342.

http://dx.doi.org/10.1037/a0025673

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

Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and

looking good: Facial cues to intelligence and their origins. *Personality and Social

Psychology Bulletin*, *28*, 238–249. http://dx.doi.org/10.1177/0146167202282009

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. http://dx.doi.org/10.1037/0022-3514.64.3.453

Züll, C. (2016). The Coding of Occupations. *GESIS Survey Guidelines*. Mannheim, Germany:

GESIS – Leibniz Institute for the Social Sciences. http://dx.doi.org/10.15465/gesis-

sg_en_019

Table 1

Intraclass Correlation Coefficients of Adult's (Study 1) and Children's (Study 2) Ratings of

the Targets' Overall (Combined) Photo Scores

	Overall photo scores		
	ICC1	ICC2	*k*
Adults' Ratings (Study 1)			
Leadership preferences			
Preselection as coach	.20	.91	39.36
Preselection as supervisor	.21	.91	39.74
Choice as leader composite[a]	.28	.94	38.65
Trait inferences			
Competence	.19	.90	39.32
Dominance	.18	.90	39.87
Facial Maturity	.18	.91	38.85
Likeable	.21	.92	40.15
Trustworthiness	.17	.89	39.48
Strong Leadership	.20	.91	39.62
Children's Ratings (Study 2)			
Leadership ratings			
Calms sailors down	.07	.72	35.15
Chooses a good passage	.06	.70	35.11
Explains what to do	.07	.73	35.13
Looks like a captain	.08	.76	35.28
Preselection as captain	.08	.74	34.18
Leadership composite[b]	.12	.83	36.10

Note. $N = 130$ targets in each condition.

ICC1 = Single-score intraclass correlation coefficient, ICC2 = Average-score intraclass

correlation coefficient; k = Average number of raters contributing to the mean for each target.

[a] An average score consisting of the two leadership ratings listed above.

[b] An average score consisting of the five leadership ratings listed above.

Table 2

Factor Loadings and Variance Explained for Principal Components Analyses in Study 1

Trait Inferences	Warmth	Power
Dominance	.14	**.94**
Facial Maturity	.09	**.86**
Likeable	**.97**	.09
Trustworthiness	**.95**	.18
Competence	.76	.62
Strong Leadership	.53	.81
Variance explained	45.4%	44.8%

Table 3

Means, Standard Deviations, and Correlations of Variables in Study 1

Variable	*M*	*SD*	1	2	3	4	5	6	7
1. Choice as Leader	2.31	0.39							
2. Warmth	3.85	0.71	.90**						
3. Power	4.26	0.63	.54**	.27**					
4. Age	46.00	7.81	.04	-.14	.13				
5. Professional Coach status	0.08	.27	.12	.07	.12	.06			
6. Core self-evaluations	3.83	0.42	.16	.17	.03	.12	.11		
7. Occupational status	7.17	2.88	.24	.18	.13	.18*	.24**	.30**	
8. Leadership role	2.65	1.39	.20	.17	.02	.18*	.16	.22*	.49**

Note. $N = 130$.

We used a cluster-robust variance estimator in the predictive models and do not report significance levels for the relations between appearance and age, Professional Coach status, core self-evaluations, occupational status, and leader role because these correlations do not correct for the clustering by photoset (Antonakis et al., 2010).

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$.

Table 4

Unstandardized Regression Coefficients, Corresponding Standard Error Estimates, and Test Statistics for Predicting Core Self-Evaluations, Actual Leadership Role, and Occupational Status in Studies 1and 2

	Study 1: Adults								Study 2: Children							
	Core Self-Evaluations								Core Self-Evaluations							
	Model 1a		Model 1b		Model 2a		Model 2b		Model 1a		Model 1b		Model 2a		Model 2b	
Choice as Leader/Leadership	0.17**	(.05)	0.16**	(.05)	0.16**	(.06)	0.15**	(.05)	0.34**	(.11)	0.34**	(.11)	0.30**	(.11)	0.30**	(.11)
Age					0.01*	(.003)	0.01*	(.003)					0.01*	(.004)	0.01*	(.004)
Professional Coach status			0.14*	(.06)			0.13†	(.07)			0.14	(.13)			0.13	(.13)
Model Significance: Wald χ^2 [df]	10.24** (1)		18.97** (2)		8.25* (2)		33.64** (3)		10.35** (1)		11.60** (2)		11.73** (2)		12.84** (3)	
Explained variance R^2	.025		.033		.037		.044		.074		.082		.083		.090	
Instrument Strength: Wald χ^2 [df]					8.25* (2)		10.80** (2)						11.73** (2)		11.26** (2)	
LR Test vs. Cons Dist: χ^2 [df]													6.40* (2)		6.02* (2)	

	Actual Leadership Role								Actual Leadership Role							
	Model 3a		Model 3b		Model 4a		Model 4b		Model 3a		Model 3b		Model 4a		Model 4b	
Choice as Leader/Leadership	0.72**	(.27)	0.67*	(.27)					0.81*	(.36)	0.78*	(.36)				
Core Self-Evaluations					4.61*	(2.01)	4.60*	(2.05)					2.99*	(1.23)	2.92*	(1.24)
Professional Coach status			0.69**	(.14)			0.05	(.24)			.75†	(.44)			0.33	(.57)
Model Significance: Wald χ^2 [df]	7.42** (1)		27.57** (2)		5.29* (1)		9.30** (2)		5.06* (1)		8.10* (2)		5.94* (1)		7.81* (2)	
Explained variance R^2	.041		.059		.050		.050		.037		.059		.050		.054	
Disturbance correlation					-0.67*		-0.67*						-0.39†		-0.38†	
Indirect effect via CSEs					0.74**		0.69**						0.91**		0.87**	

	Occupational Status				Occupational Status			
	Model 5a	Model 5b	Model 6a	Model 6b	Model 5a	Model 5b	Model 6a	Model 6b

—

Choice as Leader/Leadership	1.80**	(.42)	1.62**	(.37)					1.54*	(.75)	1.43†	(.73)			
Core Self-Evaluations					10.57*	(4.35)	10.21*	(4.26)					5.82*	(2.37)	5.50*	(2.35)
Professional Coach status			2.25*	(.95)			0.85	(.86)			2.43**	(.91)			1.62	(1.08)
Model Significance: Wald χ^2 [df]	18.58** (1)		20.37** (2)		5.91* (1)		7.49* (2)		4.21* (1)		11.63** (2)		6.04* (1)		11.78** (2)	
Explained variance R^2	.059		.102		.090		.097		.031		.082		.090		.111	
Disturbance correlation					-1.47*		-1.42*						-0.65		-0.61	
Indirect effect via CSEs					1.71**		1.54**						1.76**		1.65**	

Note. $N = 130$ targets. Standard error estimates in parentheses. Models estimated with cluster robust standard errors (due to photoset nesting) using maximum likelihood in Study 1. We used an instrumental variable estimator to obtain consistent estimates for core self-evaluations predicting the outcomes (Models 4a, 4b, 6a, 6b) by using choice as leader (Study 1) or leadership (Study 2) and age as instruments. In Models 4a, 4b, 6a, and 6b, we calculated R^2 by squaring the correlation of y with its predicted value $\hat{y}$ (Antonakis et al., 2010). CSEs: Core self-evaluations. LR Test vs. Cons Dist: Likelihood ratio test vs. a model with constrained disturbances. † $p < .10$, * $p < .05$, ** $p < .01$ (or 99% $CI > 0$).

Table 5

Factor Loadings and Variance Explained for Principal Components Analysis in Study 2

Trait Inferences	Leadership factor
Calms sailors down	.92
Chooses a good passage	.91
Explains what to do	.95
Looks like a captain	.87
Preselection as captain	.90
Variance explained	82.6%

Table 6

Means, Standard Deviations, and Correlations of all Variables in Study 2

Variable	*M*	*SD*	1	2	3	4	5
1. Leadership	3.03	0.33					
2. Age	46.00	7.81	.02				
3. Professional Coach status	0.08	0.27	.06	.06			
4. Core self-evaluations	3.83	0.42	.27**	.12	.11		
5. Occupational status	7.17	2.88	.18*	.18*	.24**	.30**	
6. Leadership role	2.65	1.39	.19*	.18*	.16	.22*	.49**

Note. $N = 130$ targets.

* $p \leq .05$, ** $p \leq .01$.

Appendix A

The Core Self Evaluations Scale (CSES)

Below are several statements about you with which you may agree or disagree. Please indicate for every item the amount of agreement, which describes yourself best.

1 = *Strongly disagree*; 2 = *Disagree*; 3 = *Neither agree nor disagree*; 4 = *Agree*; 5 = *Strongly agree*

1. I am confident I get the success I deserve in life.

2. Sometimes I feel depressed. (r)

3. When I try, I generally succeed.

4. Sometimes when I fail I feel worthless. (r)

5. I complete tasks successfully.

6. Sometimes, I do not feel in control of my work. (r)

7. Overall, I am satisfied with myself.

8. I am filled with doubts about my competence. (r)

9. I determine what will happen in my life.

10. I do not feel in control of my success in my career. (r)

11. I am capable of coping with most of my problems.

12. There are times when things look pretty bleak and hopeless to me. (r)

r = reverse-scored.

Appendix B

The individuals serving as targets completed the following additional measures: self-monitoring (18 items; Snyder, 1974), political skill inventory (18 items; Ferris et al., 2005), positive and negative affect schedule (20 items; Watson et al., 1988), and subscales from the work values inventory (18 items; Super, 1973). They also answered demographic questions about their parents' country of origin, height, weight, education, full- versus part-time work, name of occupation, job tenure in years, hours of work per week, number of employees supervised, and prior leadership roles (i.e., which type, how long, and number of employees supervised). They furthermore responded to relevant questions related to handball: the name and league of their sport club, average age class and sex of their team, number of players on their team, their team's standing in the last season, their number of hours of training per week, and their total hours busy with handball per week. Target coaches also answered questions about their coaching tenure (overall and in their current club) and highest coaching league (and for how long). Target players answered whether they had been a coach before (and, if so, for which team in what league). Finally, the target individuals indicated whether they enacted a leadership role in their spare time in another domain (e.g., another sports club, religious organization, music ensemble, garden plot, political party, etc.) and, if so, the type of position, hours per week spent in that activity, and the number of persons they supervised.

Supplemental Materials

Results Decomposed for the Two Photo Types

Study 1. We used the same estimation strategy as described in Study 1 of the manuscript for the supplemental analyses (see Table S3 for descriptive statistics and correlations between all variables, and Table S1 for all ICCs). Consistent with Hypothesis 1, cluster-robust regressions showed that the targets' Choice as Leader scores positively related to their actual leadership role in the neutral (Table S4, Model 3a), and impression-management conditions (Table S5, Model 3a), regardless of whether the Professional Coach dummy variable was included (Table S4, Model 3b and Table S5, Model 3b). We found the same pattern of results for the relationship between Choice as Leader and occupational status in both conditions regardless of whether Professional Coach status was considered: neutral (Table S4, Models 5a and 5b), and impression management (Table S5, Models 5a and 5b). Both of these findings supported Hypothesis 1. Moreover, regardless of whether Professional Coach status was considered, Choice as Leader also positively related to CSEs in the neutral (Table S4, Models 1a and 1b), and impression-management condition (Table S5, Models 1a and 1b), supporting Hypothesis 2.

Table S4 and S5 present the summary of the path estimates for the instrumental variable regression models (Models 2a, 2b, 4a, 4b, 6a, 6b in each table). Cluster-robust mediation analyses testing Hypothesis 3 revealed that CSEs mediated the positive associations between the Choice as Leader judgments and both success criteria for each photo type: actual leadership role (neutral photo: indirect effect = 0.79, SE = 0.23, 95% CI [0.33, 1.24], Table S4, Models 2a and 4a; impression-management photo: indirect effect = 0.53, SE = 0.13, 95% CI [0.27, 0.79], Table S5, Models 2a and 4a) and occupational status (neutral photo: indirect effect = 1.84, SE = 0.41, 95% CI [1.04, 2.64], Table S4, Models 2a and 6a; impression-management photo: indirect effect = 1.17, SE = 0.35, 95% CI [0.49, 1.86], Table

S5, Models 2a and 6a) regardless of whether Professional Coach status was considered in the models (neutral photo: leadership role indirect effect = 0.73, SE = 0.24, 95% CI [0.26, 1.20], Table S4, Models 2b and 4b; occupational-status indirect effect = 1.66, SE = 0.39, 95% CI [0.90, 2.43], Table S4, Models 2b and 6b; impression management: leadership-role indirect effect = 0.50, SE = 0.15, 95% CI [0.21, 0.78], Table S5, Models 2b and 4b; occupational-status indirect effect = 1.06, SE = 0.36, 95% CI [0.36, 1.76], Table S5, Models 2b and 6b).

In addition, both correlations between the disturbances of CSEs with leadership role (Ψ = −0.83, SE = 0.42, p = .050) and with occupational status (Ψ = −1.89, SE = 0.61, p = .002) were significant for the neutral photos, and the disturbances approached significance for the impression management photos (CSEs-leadership role: Ψ = −0.57, SE = 0.29, p = .054; CSEs-occupational status: Ψ = −1.18, SE = 0.71, p = .097). When controlling for Professional Coach status, we found a similar pattern for the disturbances (neutral photos: CSEs-leadership role: Ψ = −0.84, SE = 0.44, p = .058; CSEs-occupational status: Ψ = −1.85, SE = 0.61, p = .003; impression management: CSEs-leadership role: Ψ = −0.56, SE = 0.29, p = .050; CSEs-occupational status: Ψ = −1.12, SE = 0.67, p = .094). We therefore conducted stronger tests for endogeneity and compared the likelihood ratio of each instrumental variable model to one in which we constrained the disturbances of CSEs-leadership role and CSEs-occupational status to zero. The likelihood ratio tests were significant for both photo types (neutral photos: $\chi^2(2)$ = 20.26, p < .001; impression management: $\chi^2(2)$ = 15.80, p < .001), regardless of whether Professional Coach status was considered (neutral photos: $\chi^2(2)$ = 17.49, p < .001; impression management: $\chi^2(2)$ = 12.10, p = .002), supporting the endogeneity of the CSEs and the need for an IV-estimator to obtain consistent estimates. Overidentification tests indicated that the mediation models fit the data well [neutral photos: $\chi^2(2)$ = 0.23, p = .89; with Professional Coach dummy: $\chi^2(2)$ = 0.21; p = .90; impression management photos: $\chi^2(2)$ = 0.47, p = .79; with Professional Coach dummy: $\chi^2(2)$ = 0.49; p = .78]. The modification

indices of the mediation models showed that none was above the minimum value, suggesting that our models fit the data well.

Table S2 presents the factor analyses that produced the Warmth and Power composites for both photo types. The Choice as Leader judgments strongly associated with the Warmth ratings [neutral photo: $r(128) = .88$, $p < .001$; impression-management photo: $r(128) = .88$, $p < .001$], and moderately to strongly related to Power ratings [neutral photo: $r(128) = .40$, $p < .001$; impression-management photo: $r(128) = .53$, $p < .001$]. Cluster-robust regressions additionally showed that the targets' Warmth scores positively related to their occupational status (neutral photo: $B = 0.72$, $SE = 0.22$, $p = .001$; impression-management photo: $B = 0.47$, $SE = 0.22$, $p = .03$), and CSEs (neutral photo: $B = 0.05$, $SE = 0.02$, $p = .03$; impression-management photo: $B = 0.11$, $SE = 0.04$, $p = .004$), but they only marginally related to actual leadership role (neutral photo: $B = 0.34$, $SE = 0.18$, $p = .06$; impression-management photo: $B = 0.20$, $SE = 0.10$, $p = .051$). Power did not relate to targets' leadership role, occupational status, or CSEs for either photo type, however (all $|B|$s < 0.54, all $|SE|$s < 0.51, all ps $> .29$). Mediation analyses with an instrumental variable estimator revealed that CSEs mediated the positive associations between Warmth and both success criteria: actual leadership role (neutral photo: indirect effect $= 0.42$, $SE = 0.12$, 95% CI [0.19, 0.66]; impression-management photo: indirect effect $= 0.27$, $SE = 0.08$, 95% CI [0.11, 0.42]) and occupational status (neutral photo: indirect effect $= 0.88$, $SE = 0.16$, 95% CI [0.56, 1.20]; impression-management photo: indirect effect $= 0.58$, $SE = 0.15$, 95% CI [0.29, 0.86]). Yet, CSEs did not mediate the association between Power and either success criterion for either photo type (all $|$indirect effect$|$s < 0.33, $|SE|$s < 0.50, 95% CIs [-0.17, 0.24]).

Study 2. We used the same estimation strategy described in Study 2 of the manuscript. Table S6 presents the results of the factor analysis used to produce the Leadership composite for both photo types, and Table S7 shows the descriptive statistics and correlations between all variables. The associations between Leadership and the targets' leadership roles ($B = 0.68$,

$SE = 0.34$, $p = .046$; Table S4, Model 3a) and occupational status ($B = 1.86$, $SE = 0.70$, $p = .007$; Table S4, Model 5a) were positive and significant for the neutral photos. The association between Leadership and occupational status remained significant when controlling for professional coach status ($B = 1.79$, $SE = 0.68$, $p = .008$; Table S5, Model 5b), but the association with leadership role became marginally significant ($B = 0.65$, $SE = 0.34$, $p = .051$; Table S5, Model 3b). In sum, Hypothesis 1 was mostly supported for the neutral photos.

The association between Leadership and leadership role was positive for the impression management photos ($B = 0.62$, $SE = 0.31$, $p = .046$; Table S5, Model 3a), but became marginally significant when accounting for Professional Coach status. The association between Leadership with occupational status was not significant, however (Table S5, Model 5a and 5b). Thus, Hypothesis 1 was only supported for actual leadership role for the impression-management photos when not accounting for Professional Coach status.

Consistent with Hypothesis 2, Leadership positively associated with CSEs for both photo types (neutral: $B = 0.29$, $SE = 0.10$, $p = .004$; Table S4, Model 1a; impression management: $B = 0.26$, $SE = 0.09$, $p = .005$, Table S5, Model 1a), regardless of whether Professional Coach status was considered ($B = 0.29$, $SE = 0.10$, $p = .004$; Table S4, Model 1b; impression management: $B = 0.25$, $SE = 0.09$, $p = .007$, Table S5, Model 1b).

Mediation analyses with an instrumental variable estimator revealed that CSEs again mediated the positive association between Leadership and the two success criteria for both photo types: neutral photo: leadership-role indirect effect $= 0.77$, $SE = 0.28$, 95% CI [0.23, 1.35], Table S4, Models 2a and 4a; occupational status indirect effect $= 1.89$, $SE = 0.66$, 95% CI [0.64, 3.25], Table S4, Models 2a and 6a; impression management: leadership-role indirect effect $= 0.71$, $SE = 0.27$, 95% CI [0.14, 1.20], Table S5, Models 2a and 4a; occupational-status indirect effect $= 1.08$, $SE = 0.59$, 95% CI [-0.09, 2.21], 90% CI [0.22, 2.02], Table S5, Models 2a and 6a. The mediation effect for occupational status was only marginally

significant for the impression-management photos, however. When controlling for

Professional Coach status, we obtained the same pattern of results: neutral photo: leadership-

role indirect effect = 0.75, *SE* = 0.28, 95% CI [0.19, 1.31], Table S4, Models 2b and 4b;

occupational-status indirect effect = 1.82, *SE* = 0.66, 95% CI [0.56, 3.16], Table S4, Models

2b and 6b; impression management: leadership-role indirect effect = 0.68, *SE* = 0.27, 95% CI

[0.11, 1.18], Table S5, Models 2b and 4b; occupational-status indirect effect = 0.97, *SE* =

0.58, 95% CI [-0.26, 2.07], 90% CI [0.03, 1.91], Table S5, Models 2b and 6b. In sum,

Hypothesis 3 was mostly supported.

Moreover, three of the four correlations between the disturbances of CSEs and the

outcome variables approached significance (neutral: CSEs-leadership role: $\Psi = -0.39$, *SE* =

0.23, *p* = .08; CSEs-occupational status $\Psi = -0.92$, *SE* = 0.48, *p* = .054; impression

management: CSEs-leadership role: $\Psi = -0.42$, *SE* = 0.24, *p* = .08; CSEs-occupational status

$\Psi = -0.48$, *SE* = 0.44, *p* = .27), regardless of whether Professional Coach status was

considered (neutral: CSEs-leadership role: $\Psi = -0.39$, *SE* = 0.23, *p* = .09; CSEs-occupational

status $\Psi = -0.89$, *SE* = 0.47, *p* = .06; impression management: CSEs-leadership role: $\Psi =$

-0.41, *SE* = 0.24, *p* = .09; CSEs-occupational status $\Psi = -0.42$, *SE* = 0.43, *p* = .33), indicating

that the CSEs may be endogenous.

We therefore conducted stronger tests for endogeneity and compared the likelihood

ratio of each instrumental variable model to one in which we constrained the disturbances of

CSEs-leadership role and CSEs-occupational status to zero. The likelihood ratio tests were

significant for the neutral photo type, $\chi^2(2) = 7.78$, *p* = .02 [when controlling for Professional

Coach status: $\chi^2(2) = 7.57$, *p* = .02], and approached significance for the impression-

management photos, $\chi^2(2) = 5.11$, *p* =.077 [when controlling for Professional Coach status:

$\chi^2(2) = 4.72$, *p* = .09], supporting the endogeneity of the CSEs and the need for an

instrumental variable estimator to obtain consistent estimates.[7] Overidentification tests

indicated that the mediation models fit the data well [neutral photos: $\chi^2(2) = 0.85$, $p = .65$;

with Professional Coach dummy: $\chi^2(2) = 0.81$; $p = .67$; impression management photos: $\chi^2(2)$

$= 1.82$, $p = .40$; with Professional Coach dummy: $\chi^2(2) = 1.80$; $p = .41$]. The modification

indices of the mediation models showed that none was above the minimum value, suggesting

that our models fit the data well.

A pattern similar to that in Study 1 emerged for the trait inferences: The children's

Leadership judgments strongly correlated with adults' trait inferences of Warmth [neutral

photo: $r(128) = .59$, $p < .001$; impression-management photo: $r(128) = .65$, $p < .001$], and

moderately with their judgments of Power [neutral photo: $r(128) = .32$, $p < .001$; impression-

management photo: $r(128) = .43$, $p < .001$].

---

[7] When using mediation analyses in which the disturbance of CSEs-occupational status is

constrained to zero, then the indirect effect of Leadership on occupational status via CSEs is

significant for the impression management photos (indirect effect $= 0.53$, $SE = 0.25$, 95% CI

[.14, 1.14]; when controlling for Professional Coach status: indirect effect $= 0.48$, $SE = 0.24$,

95% CI [.11, 1.06]).

Table S1

Intraclass Correlation Coefficients of Adult's (Study 1) and Children's (Study 2) Ratings of the Targets' Neutral Photo, and Impression-Management Photo Scores

	Neutral Photo			Impression Management Photo		
	ICC1	ICC2	k	ICC1	ICC2	k
Adults' Ratings (Study 1)						
Leadership preferences						
Preselection as coach	.19	.82	19.29	.24	.86	20.07
Preselection as supervisor	.21	.84	19.52	.25	.87	20.22
Choice as leader composite[a]	.27	.88	19.29	.33	.91	19.58
Trait inferences						
Competence	.19	.81	19.02	.22	.85	20.29
Dominance	.18	.81	19.18	.24	.87	20.68
Facial Maturity	.17	.80	19.03	.22	.85	19.82
Likeable	.24	.86	19.52	.28	.89	20.62
Trustworthiness	.20	.83	19.12	.19	.82	20.36
Strong Leadership	.20	.83	19.15	.23	.86	20.47
Children's Ratings (Study 2)						
Leadership ratings						
Calms sailors down	.07	.62	20.76	.08	.55	14.38
Chooses a good passage	.07	.62	20.77	.07	.52	14.34
Explains what to do	.08	.65	20.88	.05	.44	14.25
Looks like a captain	.08	.62	21.03	.10	.61	14.25
Preselection as captain	.09	.68	20.28	.08	.56	13.90
Leadership composite[b]	.13	.76	21.44	.12	.67	14.66

Note. $N = 130$ targets in each condition.

ICC1 = Single-score intraclass correlation coefficient, ICC2 = Average-score intraclass correlation coefficient; k = Average number of raters contributing to the mean for each target.

[a] An average score consisting of the two leadership ratings listed above.

[b] An average score consisting of the five leadership ratings listed above.

Table S2

Factor Loadings and Variance Explained for Principal Components Analyses Decomposed

for the Two Photo Types in Study 1

Trait Inferences	Neutral Photo		Impression-Management Photo	
	Warmth	Power	Warmth	Power
Competence	.79	.54	.68	.69
Dominance	.16	**.89**	-.02	**.98**
Facial Maturity	-.04	**.82**	.12	**.89**
Likeable	**.96**	.06	**.98**	.03
Trustworthiness	**.94**	.06	**.93**	.25
Strong Leadership	.56	.76	.44	.86
Variance explained	46.3%	39.0%	41.6%	49.7%

Table S3

Means, Standard Deviations, and Correlations for All Variables Decomposed for the Two Photo Types in Study 1

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11
Neutral Photos													
1. Choice as Leader	2.29	0.40	-	.88**	.40**	.73**	.59**	.44**	.02	.13	.13	.27	.22
2. Warmth	3.79	0.78		-	.16†	.65**	.60**	.26**	-.21	.08	.09	.19	.19
3. Power	4.24	0.63			-	.44**	.24**	.65**	.06	.11	.01	.08	-.05
Impression-Management Photos													
4. Choice as Leader	2.34	0.43				-	.88**	.53**	.05	.09	.17	.18	.16
5. Warmth	3.91	0.81					-	.22*	-.04	.05	.20	.13	.11
6. Power	4.29	0.76						-	.17	.10	.04	.14	.07
Target Characteristics													
7. Age	46.00	7.81							-	.06	.12	.18*	.18*
8. Professional Coach status	0.08	.27								-	.11	.24**	.16†
9. Core Self-Evaluations	3.83	0.42									-	.30**	.22*
10. Occupational status	7.17	2.88										-	.49**
11. Leadership role	2.65	1.39											-

Note. N = 130.

We used a cluster-robust variance estimator in the predictive models and do not report significance levels for the associations between appearance and age, Professional Coach status, core self-evaluations, occupational status, and leadership role because these correlations do not correct for the clustering by photoset (see Antonakis et al., 2010).

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$.

Table S4

Unstandardized Regression Coefficients, Corresponding Standard Error Estimates, and Test Statistics for Predicting Core Self-Evaluations, Actual Leadership Role, and Occupational Status for the Neutral Photos in Studies 1 and 2

Core Self-Evaluations

	Study 1: Adults								Study 2: Children							
	Model 1a		Model 1b		Model 2a		Model 2b		Model 1a		Model 1b		Model 2a		Model 2b	
Choice as Leader/Leadership	0.13*	(.05)	0.12*	(.06)	0.14**	(.06)	0.13**	(.05)	0.29**	(.10)	0.29**	(.10)	0.26**	(.10)	0.26**	(.10)
Age					0.01*	(.002)	0.01*	(.002)					0.01*	(.004)	0.01*	(.003)
Professional Coach status			0.14*	(.07)			0.13†	(.07)			0.15	(.13)			0.14	(.13)
Model Significance: Wald χ² [df]	6.45* [1]		21.35** [2]		6.50* [2]		137.49** [3]		8.49** [1]		9.87** [2]		9.87** [2]		11.12* [3]	
Explained variance R²	.017		.025		.029		.036		.061		.071		.071		.079	
Instrument Strength: Wald χ² [df]					6.50* [2]		5.99† [2]						9.87** [2]		9.56** [2]	
LR Test vs. Cons Dist: χ² [df]					20.26** [2]		17.49** [2]						7.78* [2]		7.57* [2]	

Actual Leadership Role

	Study 1: Adults								Study 2: Children							
	Model 3a		Model 3b		Model 4a		Model 4b		Model 3a		Model 3b		Model 4a		Model 4b	
Choice as Leader/Leadership	0.76*	(.34)	0.70*	(.34)					0.68*	(.34)	0.65†	(.34)				
Core Self-Evaluations					5.53*	(2.67)	5.59*	(2.81)					3.02*	(1.34)	2.94*	(1.35)
Professional Coach status			0.67**	(.14)			-0.11	(.30)			.78†	(.44)			0.32	(.58)
Model Significance: Wald χ² [df]	5.11* [1]		31.32** [2]		4.29* [1]		6.34** [2]		3.98* [1]		7.15* [2]		5.10* [1]		7.06* [2]	
Explained variance R²	.049		.066		.050		.049		.030		.052		.050		.054	
Disturbance correlation					-0.83*		-0.84†						-0.39†		-0.39†	

	Occupational Status				Occupational Status			
Indirect effect via CSEs			0.79**	0.73**			0.77**	0.75**
	Model 5a	Model 5b	Model 6a	Model 6b	Model 5a	Model 5b	Model 6a	Model 6b
Choice as Leader/Leadership	1.93** (.45)	1.74** (.41)			1.86** (.70)	1.79** (.68)		
Core Self-Evaluations			12.96** (4.24)	12.69** (4.30)			7.37** (2.82)	7.13* (2.81)
Professional Coach status		2.18* (.97)		0.44 (1.04)		2.44** (.89)		1.35 (1.22)
Model Significance: Wald χ^2 [df]	18.45** [1]	32.13** [2]	9.32** [1]	9.65** [2]	7.15* [1]	15.01** [2]	6.86** [1]	11.46** [2]
Explained variance R^2	.074	.114	.090	.093	.052	.104	.090	.104
Disturbance correlation			-1.89**	-1.85**			-0.92[†]	-0.89[†]
Indirect effect via CSEs			1.84**	1.66**			1.89**	1.82**

Note. $N = 130$ targets. Standard error estimates in parentheses. Models estimated with cluster robust standard errors (due to photoset nesting) using maximum likelihood in Study 1. We used an instrumental variable estimator to obtain consistent estimates for core self-evaluations predicting the outcomes (Models 4a, 4b, 6a, 6b) by using choice as leader (Study 1) or leadership (Study 2) and age as instruments. In Models 4a, 4b, 6a, and 6b, we calculated R^2 by squaring the correlation of y with its predicted value $\hat{y}$ (Antonakis et al., 2010). CSEs: Core self-evaluations. LR Test vs. Cons Dist: Likelihood ratio test vs. a model with constrained disturbances. [†] $p < .10$, * $p < .05$, ** $p < .01$ (or 99% $CI > 0$).

Table S5

Unstandardized Regression Coefficients, Corresponding Standard Error Estimates, and Test Statistics for Predicting Core Self-Evaluations, Actual Leadership Role, and Occupational Status for the Impression Management Photos in Studies 1 and 2

	Study 1: Adults								Study 2: Children							
	Core Self-Evaluations								Core Self-Evaluations							
	Model 1a		Model 1b		Model 2a		Model 2b		Model 1a		Model 1b		Model 2a		Model 2b	
Choice as Leader/Leadership	0.16**	(.06)	0.15**	(.06)	0.13*	(.06)	0.13*	(.05)	0.26**	(.09)	0.25**	(.09)	0.22*	(.09)	0.22*	(.09)
Age					0.01*	(.003)	0.01*	(.003)					0.01*	(.004)	0.01*	(.004)
Professional Coach status			0.14*	(.06)			0.13†	(.07)			0.14	(.13)			0.13	(.13)
Model Significance: Wald χ^2 [df]	6.67** [1]		10.71** [2]		5.74† [2]		6.64† [3]		7.72** [1]		8.93* [2]		9.28** [2]		10.37* [3]	
Explained variance R^2	.027		.035		.038		.045		.056		.064		.067		.074	
Instrument Strength: Wald χ^2 [df]					5.74† [2]		5.85† [2]						9.28** [2]		8.81* [2]	
LR Test vs. Cons Dist: χ^2 [df]					15.80** [2]		12.10** [2]						5.11† [2]		4.72† [2]	

	Actual Leadership Role								Actual Leadership Role							
	Model 3a		Model 3b		Model 4a		Model 4b		Model 3a		Model 3b		Model 4a		Model 4b	
Choice as Leader/Leadership	0.50**	(.17)	0.46*	(.19)					0.62*	(.31)	0.59†	(.31)				
Core Self-Evaluations					4.01*	(1.87)	3.97*	(1.86)					3.18*	(1.41)	3.09*	(1.35)
Professional Coach status			0.74**	(.14)			0.16	(.29)			.76†	(.44)			0.30	(.58)
Model Significance: Wald χ^2 [df]	8.72** [1]		33.33** [2]		4.58* [1]		12.47** [2]		3.99* [1]		6.99* [2]		5.06* [1]		6.82* [2]	
Explained variance R^2	.025		.045		.050		.051		.030		.051		.050		.053	
Disturbance correlation					-0.57†		-0.56*						-0.42†		-0.41†	

	Occupational Status								Occupational Status							
	Model 5a		Model 5b		Model 6a		Model 6b		Model 5a		Model 5b		Model 6a		Model 6b	
Indirect effect via CSEs					0.53**		0.50**						0.71**		0.68**	
Choice as Leader/Leadership	1.22*	(.50)	1.09*	(.51)					0.69	(.65)	0.59	(.63)				
Core Self-Evaluations					8.85*	(4.50)	8.47*	(4.29)					4.82†	(2.58)	4.38†	(2.57)
Professional Coach status			2.37**	(.90)			1.13	(.76)			2.47**	(.91)			1.80†	(1.03)
Model Significance: Wald χ^2 [df]	5.86* [1]		7.64* [2]		3.88* [1]		7.02* [2]		1.14 [1]		8.50* [2]		3.50† [1]		10.13** [2]	
Explained variance R^2	.034		.082		.090		.100		.009		.061		.090		.117	
Disturbance correlation					-1.18†		-1.12†						-0.48		-0.42	
Indirect effect via CSEs					1.17**		1.06**						1.07*		0.97†	

Note. $N = 130$ targets. Standard error estimates in parentheses. Models estimated with cluster robust standard errors (due to photoset nesting) using maximum likelihood in Study 1. We used an instrumental variable estimator to obtain consistent estimates for core self-evaluations predicting the outcomes (Models 4a, 4b, 6a, 6b) by using choice as leader (Study 1) or leadership (Study 2) and age as instruments. In Models 4a, 4b, 6a, and 6b, we calculated R^2 by squaring the correlation of y with its predicted value $\hat{y}$ (Antonakis et al., 2010). CSEs: Core self-evaluations. LR Test vs. Cons Dist: Likelihood ratio test vs. a model with constrained disturbances. † $p < .10$ (or 90% $CI > 0$), * $p < .05$ (or 95% $CI > 0$), ** $p < .01$ (or 99% $CI > 0$).

Table S6

Factor Loadings and Variance Explained for Principal Components Analyses Decomposed

for the Two Photo Types in Study 2

Trait Inferences	Neutral Photo Leadership factor	Impression-Management Photo Leadership factor
Calms sailors down	.88	.88
Chooses a good passage	.91	.85
Explains what to do	.92	.93
Looks like a captain	.82	.84
Preselection as captain	.90	.84
Variance explained	78.6%	74.9%

Table S7

Means, Standard Deviations, and Correlations of all Variables Decomposed for the Two

Photo Types in Study 2

Variable	*M*	*SD*	1	2	3	4	5	6	7
Neutral Photos									
1. Leadership	3.02	0.35	-	.59**	.03	.04	.25**	.23*	.17*
Impression-Management Photos									
2. Leadership	3.04	0.39		-	.00	.06	.24**	.09	.17*
Target Characteristics									
3. Age	46.00	7.81			-	.06	.12	.18*	.18*
4. Professional Coach status	0.08	0.27				-	.11	.24**	.16†
5. Core self-evaluations	3.83	0.42					-	.30**	.22*
6. Occupational status	7.17	2.88						-	.49**
7. Leadership role	2.65	1.39							-

Note. N = 130.

† *p* ≤ .10, * *p* ≤ .05, ** *p* ≤ .01.