We Look Like Our Names: The Manifestation of Name Stereotypes in Facial Appearance

Yonat Zwebner
The Hebrew University of Jerusalem

Anne-Laure Sellier
HEC Paris

Nir Rosenfeld
The Hebrew University of Jerusalem

Jacob Goldenberg
Interdisciplinary Center (IDC) and Columbia University

Ruth Mayo
The Hebrew University of Jerusalem

Research demonstrates that facial appearance affects social perceptions. The current research investigates the reverse possibility: Can social perceptions influence facial appearance? We examine a social tag that is associated with us early in life—our given name. The hypothesis is that name stereotypes can be manifested in facial appearance, producing a *face-name matching effect*, whereby both a social perceiver and a computer are able to accurately match a person's name to his or her face. In 8 studies we demonstrate the existence of this effect, as participants examining an unfamiliar face accurately select the person's true name from a list of several names, significantly above chance level. We replicate the effect in 2 countries and find that it extends beyond the limits of socioeconomic cues. We also find the effect using a computer-based paradigm and 94,000 faces. In our exploration of the underlying mechanism, we show that existing name stereotypes produce the effect, as its occurrence is culture-dependent. A self-fulfilling prophecy seems to be at work, as initial evidence shows that facial appearance regions that are controlled by the individual (e.g., hairstyle) are sufficient to produce the effect, and socially using one's given name is necessary to generate the effect. Together, these studies suggest that facial appearance represents social expectations of how a person with a specific name should look. In this way a social tag may influence one's facial appearance.

Keywords: face perception, naming, self-fulfilling prophecy, social influence, stereotypes

Supplemental materials: http://dx.doi.org/10.1037/pspa0000076.supp

We are told not to judge a book by its cover, but we all do. Facial appearance matters when we judge people: It affects how a person is perceived and judged by others regarding critical characteristics such as intelligence, trustworthiness, attractiveness, warmth, dominance, and so forth (Ballew & Todorov, 2007; Bar, Neta, & Linz, 2006; Berry, 1991; Berry & Brownlow, 1989; Penton-Voak, Pound, Little, & Perrett, 2006; Rule & Ambady, 2008; Todorov, 2008; Willis & Todorov, 2006).

Could it also be the other way around? Can the way people judge us affect how we look? Up to now, research has mainly focused on how social perceptions are influenced by facial appearance (Oosterhof & Todorov, 2008; Todorov, Said, Engell, & Oosterhof, 2008; Walker & Vetter, 2016) rather than whether facial appearance can be influenced by social perceptions. We examine this reverse possibility and suggest that the association between faces and social perceptions could be a two-way street.

This article was published Online First February 27, 2017.

Yonat Zwebner, School of Business Administration, The Hebrew University of Jerusalem; Anne-Laure Sellier, Department of Marketing, HEC Paris; Nir Rosenfeld, The Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem; Jacob Goldenberg, Arison School of Business, Interdisciplinary Center (IDC), and Visiting Professor at Columbia Business School, Columbia University; Ruth Mayo, Department of Psychology, The Hebrew University of Jerusalem.

For interested readers, data from all experiments are available at https://osf.io/

We express our gratitude to Hillel Aviezer, Shai Danziger, Andrew D. Gershoff, Shira Goldenberg, and David Mazursky for insightful comments on previous versions of this paper and project. Finally, we thank Charles M. Judd, Jacob Westfall, and Ariel Goldstein for their advice regarding the mixed model analysis, and Noam Siegelman for his comprehensive assistance with conducting the mixed model analysis. We also thank the HEC Foundation for funding part of this research. This research is part of Yonat Zwebner's PhD thesis.

Correspondence concerning this article should be addressed to Yonat Zwebner, School of Business Administration, The Hebrew University of Jerusalem, Jerusalem, Israel. E-mail: yonat.zwebner@mail.huji.ac.il

Journal of Personality and Social Psychology, 2017, Vol. 112, No. 4, 527–554

0022-3514/17/$12.00 http://dx.doi.org/10.1037/pspa0000076

Focusing on the case of name stereotypes, we investigate whether a social tag can influence our facial appearance. In eight studies we test whether a perceiver is able to choose among a set of few names the correct given name of a person significantly above chance level according to facial appearance alone. These studies further control for socioeconomic explanations, eliminate alternative explanations and find the effect using a computer-based paradigm. Figure 1 shows an example of a trial choice from Study 5. Upon looking at the face in the figure and considering the four options given (Jacob, Dan, Josef, Nathaniel), participants chose the name "Dan" 38% of the time, which is above the 25% chance level of a random guess. This choice was correct: Dan is the real name of the person in the picture, suggesting that a person's given name may be manifested in her or his facial appearance. Different from other social factors such as gender or ethnicity, a name is an external social factor, which is chosen and not innate, therefore representing an ultimate social tag. The demonstration of our name being manifested in our facial appearance illustrates the great power that a social factor can have on our identity, potentially influencing even the way we look. Next we elaborate on the theoretical underpinnings leading to our prediction.

Conceptual Background

The Dorian Gray Effect

A name has a social meaning and can impart significant and differentiated impressions, both positive and negative (Mehrabian, 1997). Individuals attribute characteristics to people based on their given name (Mehrabian, 2001), and this name can lead to social consequences (Fryer & Levitt, 2004; Laham, Koval, & Alter, 2012). That is, names may carry personal significance, thus making them a potential contributor to the development of a person's identity, behaviors, and choices (e.g., Flugel, 1930; Insaf, 2002; Oberndorf, 1920; Shoval, Manor, Nahshoni, Weizman, & Zalsman, 2012). It has been shown that appearance can be affected by identity, choices, and behaviors (Penton-Voak et al., 2006; Walker & Vetter, 2016); thus we suggest that a name might eventually affect physical appearance. One's identity, choices, and behaviors may have as direct an effect on appearance as choosing a hairstyle, but these factors can also indirectly affect the choice of our lifestyle, work, temperament, and so forth, all of which could eventually affect our facial appearance. Moreover, as we expand on later, names carry their own stereotypical facial imagery; these face-name stereotypes may of course also affect one's appearance and related choices.

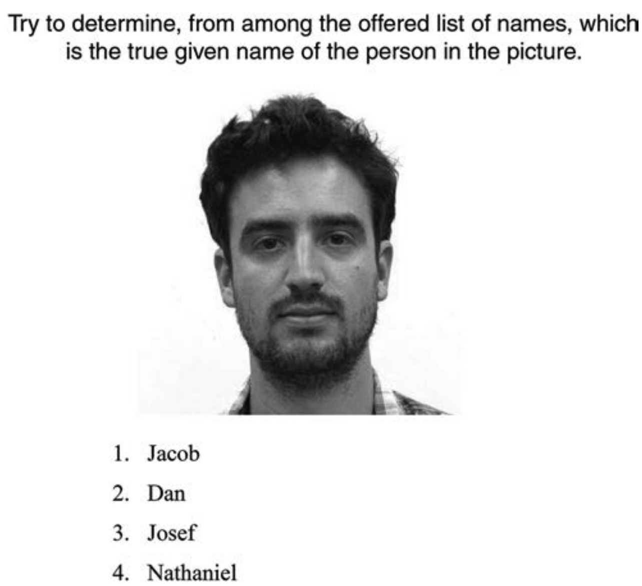

Figure 1. Example of a trial in our studies. For randomization purposes, the name Dan appeared with several different filler names for different participants.

When internal factors such as personality influence external facial appearance, this has been termed "the Dorian Gray effect" in prior research (e.g., Zebrowitz & Collins, 1997), after Oscar Wilde's (1890/1974) novel in which the deeds of the protagonist affected his portrait. We propose that one's given name may have a Dorian Gray effect on one's face. Our given name is our very first social tagging. Each name has associated characteristics, behaviors, and a look, and as such, it has a meaning and a shared schema within a society. These name stereotypes include a prototypical facial appearance such that we have a shared representation for the "right" look associated with each name. Over time, these stereotypical expectations of how we should look may eventually manifest in our facial appearance. If so, a perceiver (i.e., one who is asked to match the name to the face), as well as a computer program, should be able to accurately match a person's true name to his or her face, significantly above chance level and beyond socioeconomic factors (such as age and ethnicity). Such matching abilities might indicate the influence of a social tag (i.e., one's name) on facial appearance.

The leading conceptualization in social–cognitive psychology is that facial appearance affects how people are perceived, which in turn affects their personality. As we are social creatures, these impressions that others have of us based on our facial appearance might affect our personality by either self-fulfilling or self-defeating prophecy effects, as we develop the personality that other people expect us to exhibit ([e.g., self-fulfilling prophecy] or a personality that contradicts these expectations [e.g., self-defeating prophecy], Berry, 1991; Berry & Brownlow, 1989; Penton-Voak et al., 2006; Todorov et al., 2008; Zebrowitz, 1996, 1997; Zebrowitz & Collins, 1997; Zebrowitz, Collins, & Dutta, 1998). Scant attention, if any, has been paid to the possibility that social factors such as a name tag can have a potential effect on facial appearance.

This bias may be attributable to the tendency of lay people as well as researchers to consider facial appearance as given (see this claim also in Zebrowitz, Collins, & Dutta, 1998). Changes in facial appearance are perceived to be mainly attributable to biological variables such as genes and hormone levels affecting both our look and our personality (Jones et al., 2015; Lefevre et al., 2014) and/or to external events (such as surgeries, accidents, etc.). However, to the best of our knowledge only few studies looked at the possible influence social factors may have on our facial appearance. These studies have already demonstrated a possible reverse influence of various factors on facial appearance, supporting the Dorian Gray effect: For example, short-tempered people were found to tense specific facial muscles more than relaxed people, leading in turn to a particular development of the jaw (Kreiborg, Jensen, Moller, & Bjork, 1978). Similarly, the neutral facial expression of people with a hostile disposition was perceived as angry (Malatesta, Fiore, & Messina, 1987). However, these studies did not rule out the possibility of a facial information sign that started this connection early in life. In other words, the chicken-and-egg problem of whether it is the personality that affects facial appearance or the other way around is difficult to resolve. A pioneering study suggesting the directional influence of personality on appearance used records of both personality and appearance of subjects from 1930 to 1990. It demonstrated a Dorian Gray effect whereby early personality produces a congruent later appearance (Zebrowitz et al., 1998). Specifically, women with more sociable personalities in adolescence improved with age in their perceived physical attractiveness. Another demonstration of the effect of one's choices and resulting inner state on facial appearance is the finding that the clothes women wear affect the perceived attractiveness of their face (Lõhmus, Sundström, & Björklund, 2009). Women were photographed while wearing clothes in which they felt attractive, unattractive, or comfortable. The clothes were not visible in the picture, and the women were asked to maintain a neutral expression in all pictures. Still, without seeing the clothes, men who looked at the pictures rated the women wearing attractive clothes as more attractive than those wearing unattractive or comfortable articles of clothing, suggesting that the situational emotion that a woman felt was expressed in her face and a social observer could perceive it. These two studies support the Dorian Gray effect and are important as a first (albeit limited) step to acknowledge that our facial appearance can be influenced. The former is limited to a conflict between one's personality and appearance at an early age and the resulting personality and appearance at a later age (Zebrowitz et al., 1998), and the latter is limited to momentary effects (the clothes that women wore—Lõhmus, Sundström, & Björklund, 2009).

One exception is a work done by Zajonc, Adelmann, Murphy, and Niedenthal (1987), suggesting that social interaction can influence facial appearance by demonstrating an increase in apparent physical resemblance of spouses. Interestingly, this resemblance was associated with greater reported marital happiness and was found only after 25 years of cohabitation. To add to these findings and to overcome the methodological difficulty in controlling for the sequence of events to establish the influence of a factor on facial appearance and not the other way around, we aim to test a possible social influence on facial appearance by using a social labeling that takes place as early as possible. Moreover, we aim to test a social labeling whose influence cannot be accounted for by possible biological explanations such as genes and hormone levels. Thus, our objective is to find a social label (or tag) given to a person before his or her facial appearance can lead to any inferences, and to see whether this tag affects his or her facial appearance. A natural candidate, as explained above, is one's given name.

Given Names and Self-Fulfilling Prophecy

In most societies, a given name is chosen for us by others, usually our parents, having little to do with our facial appearance as we have just been born. Often the name is chosen before birth. We argue that once a baby receives a name, this social label (name) leads to certain social expectations, inferences, and interactions. It is possible that he or she is treated as if certain personality traits based on his or her name are already in place. In turn, the social attitude that a name triggers and its influence on one's identity (Flugel, 1930; Insaf, 2002; Oberndorf, 1920; Shoval et al., 2012) may affect the individual's self-perception and development of traits, which eventually may be realized in his or her facial appearance (e.g., hairdo, expressions, and a specific facial feature). A self-fulfilling prophecy of a stereotype regarding people's personality is well documented (Chen & Bargh, 1997; Darley & Fazio, 1980; Shih, Pittinsky, & Ambady, 1999; Sinclair, Hardin, & Lowery, 2006). More broadly, social-structuring literature has demonstrated that people apply cultural stereotypes to their self, and that stereotypes can influence people's behavior (Crocker, Major, & Steele, 1998; Greenwald et al., 2002; Swim & Stangor, 1998; Wheeler & Petty, 2001). Therefore, in the sense of a stereotype affecting people, the current research hypothesis is almost obvious. Given the critical role of social structuring in shaping who we are, why shouldn't our facial appearance be affected by social structuring as well?

Building on the self-fulfilling prophecy literature (Zebrowitz, 1996, 2011), we posit that our facial features may change over the years to eventually represent the expectations of how we should look. The process may be of a direct nature, based on the expectation of a specific look for a specific name and/or may be governed by the mediation effect of a name on personality, as recent studies have demonstrated the manifestation of personality in facial appearance (Penton-Voak et al., 2006; Walker & Vetter, 2016). Both possibilities lead to the current research hypothesis that a name can be realized in the long run in one's facial appearance.

It is critical to note that an underlying assumption of this research is that the name given to a newborn is only minimally affected by his or her look, if at all, as newborn faces are more homogeneous in appearance compared with adult faces and very little can be inferred from an infant's facial appearance (Gordon & Tanaka, 2011; Yovel et al., 2012). This research conjecture is also supported by many cases where one's name is decided before birth. We return to this point in the General Discussion.

In sum, this research tests a specific type of Dorian Gray effect, that is, the reflection of given names on faces. We term this the *face-name matching effect*, defined as the ability of a perceiver (human or computerized) to choose the correct given name of a person, significantly above chance level, according to facial appearance alone.

Face-Name Prototypes: What Is the "Right" Name for a Face?

For a name to generate a face-name matching effect, people should have shared representations in mind regarding its social labeling. Each name is associated with a character, behaviors, and a facial appearance. Only if there are clear shared name schemas, and more specifically face-name prototypes, may the social actor alter his appearance accordingly and the social perceiver identify the name from his or her facial appearance.

What are face-name prototypes? People typically recognize individuals by their face or name. From a social perceiver's perspective, identifying a familiar face involves matching that face to previously stored identity information, including one's name (Bruce & Young, 1986; Tanaka, 2001; Valentine, Bredart, Lawson, & Ward, 1991). Interestingly, it is suggested that giving a name to a newborn enables individuation and therefore better recognition (Yovel et al., 2012). Familiar faces are referred to by name rather than other possible categorization labels (Tanaka, 2001), suggesting that names play a crucial role in the perceptual formation and retrieval of face representations (Gordon & Tanaka, 2011). Consequently, these repeated representations, associating a specific face with a specific name and vice versa might generate more general face-name prototypes.

Several factors can lead to these face-name prototypes being shared within a society, eventually generating a shared representation of the "right name" for a specific face. One possible factor can be a shape-name "fit." This has been demonstrated for objects in theories based on the *bouba-kiki* effect (Köhler, 1929), which suggests that the match between an object and a name is not completely arbitrary. Instead, object names can be associated with their physical attributes (Maglio, Rabaglia, Feder, Krehm, & Trope, 2014). This could also be true for stereotypes of peoples' names, as when people hearing a name that has a "rounder" sound (e.g., Bob) have in mind a shared representation of a "rounder face" (Lea, Thomas, Lamkin, & Bell, 2007; Sapir, 1929). Probably many other factors potentially create shared face-name schemas—for instance, the literal meaning of a name or an association with famous people. There also might be a social-inference correlation, which can explain why we share face-name representations. Converging evidence suggests that both faces and names separately convey social signals, eventually leading to a possible correlation between the two. Specifically, faces communicate information to the perceiver about a person's identity (e.g., trustworthiness or aggressiveness; Cogsdill, Todorov, Spelke, & Banaji, 2014; Hassin & Trope, 2000; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015; Zebrowitz, 2006), including personality characteristics such as trustworthiness, competence, and likability (Haselhuhn, Wong, Ormiston, Inesi, & Galinsky, 2014; Livingston & Pearce, 2009; Rule & Ambady, 2008; Todorov, Mandisodza, Goren, & Hall, 2005; Van't Wout & Sanfey, 2008; Willis & Todorov, 2006). Similarly, people infer personal characteristics from one's name (e.g., popularity, successfulness; Mehrabian, 1997, 2001). Thus, because both faces and names lead to social inferences, the possibility of an existing correlation between the two is plausible.

Overall, different affiliations for names and faces should merge to generate shared general representations, eventually enabling a shared representation regarding what the "right name" most likely is for a specific face. Indeed, research has demonstrated the existence of shared face-name prototypes (Lea et al., 2007), as people often have common stereotypes about how those with particular names should look, such that they are able to agree to some extent that a name matches a specific morphed face. Specifically, Lea et al. (2007) asked participants to create face exemplars for 15 names. Those chosen to be good exemplars were morphed to create face-name prototypes. A different set of participants naming the prototypes showed convergence for some of the names, suggesting a shared perception of what specific names *should* look like. These findings propose that face-name prototypes exist. It is important to note that this study did not use real faces and did not test for face-name accuracies. Rather, it only tested whether people agree on the prototypical face for a specific name. Given the shared face-name prototypes, our research question is whether they manifest in actuality such that the stereotypical look for a specific name will eventually be represented in the face of a person carrying that name, enabling social perceivers—and even a computer program—to recognize one's true name according to his or her face, above chance level.

Overview of Studies

The current research aims to test the possibility that name stereotypes can be manifested in one's facial appearance. Empirical support for the face-name matching effect is developed using eight studies. In six studies, we measured how well individuals can match the true name to a face, in different setups. In two additional studies, we used a computerized paradigm to test the traces of names in the faces, controlling for possible human biases of the social perceiver.

For clarity we group studies in sections that address different parts of the general research question. In Section I (Studies 1 through 4), we test for the existence of the face-name matching effect. Specifically, in Studies 1A and 1B, we show that people accurately identified a person's name based only on a headshot, beyond chance level. We used different methodologies for selecting filler names to control for possible confounds related to those filler names. In Study 2, we generalized the results by examining whether the face-name matching effect can be replicated in the independent setting of a different country and culture. In Study 3, we tested several alternative explanations and methodological issues that might account for our results. Specifically, we further controlled for possible confounds related to the names, such as the presentation order, the method used to choose them, and their frequency in the general population. Additionally, we tightly controlled for socioeconomic factors (such as age and ethnicity). In Study 4, using more than 94,000 faces and a computer-based paradigm, we provided additional evidence showing it is the face itself that contains information for correctly matching names, thus ruling out the possibility of any human-based biases. Notably, our studies entail a multimethod approach and thus tackle various alternative accounts such as the possible confounds related to filler names differing from target names. Possible differences between the filler and target names were a concern regarding the face-name matching effect. Accordingly, Studies 1B, 3, and 4 were designed to directly test this possibility.

In Section II, we begin investigating the underlying mechanism responsible for the face-name matching effect by testing the role of name stereotypes. If name stereotypes cause the face-name matching effect, then the effect should be attenuated or eliminated when there is no schema available to the social perceiver. Study 5 demonstrates the importance of existing shared schemas by showing that the face-name matching effect does attenuate when a shared schema does not exist (i.e., different cultures of the target faces and participants).

Finally, in Section III we more directly test self-fulfilling prophecy as the possible underlying mechanism. First, we demonstrate that one's name is represented even in one feature of our facial appearance, the one that is most controlled by its possessor, his or her hairstyle (Study 6). Second, in Study 7 we once again used a computer-based paradigm, this time to verify that it is the "controlled" facial regions that carry more weight in producing the face-name matching effect. By analyzing the matching accuracy of the algorithm after removing different parts of a face, we created heat maps for each name to identify the facial components that are more important for a correct match to occur. Third, we provide additional support for a self-fulfilling prophecy account by demonstrating that name usage is crucial for the development of face-name matches (Study 8): The face-name matching effect occurs when depicted targets socially use their given name (e.g., Charlotte goes by Charlotte), but attenuates for depicted targets who go by an exclusive nickname rather than their given name (e.g., Charlotte goes by Chatou).

A common procedure was used for all studies, whereby we presented the perceiver (either human participants or a computer) with a series of unfamiliar headshots. For each headshot, the perceiver was given several names to choose from, among which we included the depicted person's true name. Perceivers were asked to select the true name of the person in the picture. Figure 1 illustrates an example of a trial choice. We also measured confidence in the first studies and found no relation between confidence in choice and accuracy; therefore, we report this measure only in the supplementary materials.

Target Faces

To ensure a nonsystematic collection we used several different methodologies: In some cases, hypothesis-blind research assistants obtained target faces' photographs from the Internet (Study 1A), whereas in other cases we photographed individuals who came to the university's lab in response to announcements we posted on campus boards (Studies 1B, 2, 3, 5, 6, and 8), or we conducted a wider search when target faces were not readily available, asking individuals to email us their photos (Studies 2 and 5). All photos were posed front-on and with a neutral expression. Minimal or no cosmetics were worn. Faces were presented on a computer screen with a quality of 400 × 400 pixels. In Studies 1B, 3, 6, and 8, because we photographed all individuals, we were able to control for a standard look, size, and quality of photo. However, in Studies 1A, 2, and 5, because we did not take some of the photos, quality was not at the same consistent level. Nevertheless, our focus was on standardizing the object in the photos (such that all photos contained only a headshot and seemed taken from a similar distance), so we adjusted the size of all photos to be similar enough that an unprofessional eye would not notice the differences. The computerized studies (Studies 4 and 7) required a large collection of images. We therefore created a dataset of almost 100,000 profile images of members of a popular business-oriented social platform in France, all downloaded from the same public domain. We chose this as our source of images because its users are relatively homogenous, and the facial images are typically "clean" in terms of pose, orientation, facial expression, lighting, and so forth.

Filler Names

For Study 1A, a hypothesis-blind research assistant randomly drew filler names from known local websites containing catalogs of names. Specifically, she generated a random list of numbers according to the required number of filler names, and then drew the name that was ordered in the website in accordance with the list of numbers. If a name was unusual, it was excluded and replaced with the next name presented online. For Studies 1B and 4, for methodological purposes we used other targets' names as filler names. For Studies 3, 5 (the Israeli stimuli), and 6, we used filler names of nontarget individuals who share the same age and origin of the target faces. For Studies 2 and 5 (French stimuli), a hypothesis-blind research assistant drew filler names from the Database of Given Names (INSEE 2011). For Study 8, a hypothesis-blind research assistant randomly drew filler given names from known local websites containing catalogs of names, and made up exclusive filler nicknames. Filler names did not belong to any face that appeared in the studies (except in Studies 1B and 4, where we aimed to use other targets' names as filler names); and each name—the true and the filler—appeared only once for each participant (except in Study 1B and the computerized studies). When filler names were randomly drawn from local websites containing catalogs of names, uncommon and unusual names were excluded and replaced with other names in the catalog. For each of the human-based studies, we first generated a list of filler names and then, for randomization purposes, created a few additional lists, all containing the identical filler names (except Study 2) but in different randomized orders (according to the number of experimental versions of each study). We next integrated the names from each list with the faces from one of the experimental versions. Eventually, each experimental version contained the same filler names (per experiment) but in different randomized orders. We further detail selection and presentation procedures in the section on each study and in the supplementary materials. All studies meet the ethical guidelines and requirements of the study country.

Because we aim to explore the link between a person's name and facial appearance, and because research has demonstrated that the external features (e.g., hairstyle) of unfamiliar faces are essential for recognition (Ellis, Shepherd, & Davis, 1979), we kept targets' original hairstyles and focused on real faces with their everyday appearance (except in Study 6, where we isolated both the hairstyle and the facial features).

Section I—Testing the Face-Name Matching Effect

The main goal in Section I was to test the ability of perceivers (social or computerized) to accurately match faces and names. First, in Study 1A we examined whether people can accurately identify a person's name based only on a headshot. In Study 1B, we aimed to replicate the effect while eliminating possible confounds related to the use of external filler names (we used target names as fillers and reshuffled the stimuli).

Study 1A: Accurately Matching the True Given Name to its Face

In Study 1A we examined whether the face-name matching effect exists in real life, by asking participants to choose among five proposed names the true given name of the person depicted in the picture for 20 unfamiliar faces.

Method.

Participants. One hundred twenty-one Israeli students from a local online panel (60 women, all native speakers, $M_{age} = 27.77$ years [$SD = 4.08$]) participated in the study in return for the equivalent of US$2.30.

Materials. For target selection, a hypothesis-blind research assistant nonsystematically retrieved photos from different individuals' profile pages on a large social network. All targets were Israeli young adults. The website does not report the age of the targets nor when each photograph was taken. Therefore, we attempted to select young adults that seemed to be in the age range of 20 to 30. Because of this uncertainty regarding targets' precise ages, we control for age in the following studies, where we had more control over the photos taken. We presented 20 photos (10 female) one at a time on a computer screen. For each target, five suggested given names appeared below the headshot, including the true name of the depicted person and four filler names. The filler names did not belong to any face that appeared in the study, and all names (the 20 true names and 80 filler names) appeared only once for each participant. A hypothesis-blind research assistant randomly drew these filler names from known local websites containing catalogs of names. We generated five different versions of the study, each containing the same filler names but in a different randomized order. In each version, we randomly generated the order of the faces as well as the location of the true names.

Procedure. We randomly assigned participants to one of the five study versions and informed them that the study was about impression formation. They were told that they would see pictures of people and be asked to choose one name of five presented that they thought was the true name of the person in each picture. Specifically, the opening instructions were as follows:

> This research examines impression formation regarding people. Many times in life, we meet a person for the first time or only see his or her picture, and need to make a few decisions without any additional information. This study tests situations of this kind, and its goal is to examine the first impression people have of other individuals according to their facial appearance. Therefore, there is no right or wrong answer but only your "gut feeling." Next, photos of individuals will appear. Your task now is to determine what is the true given name of each one of them.

All participants began the study by performing the same single practice trial. Next, they were exposed to the target pictures. Testing for consistency, we included two sequential and identical rounds of judgments (i.e., participants were exposed to the same block of 20 target pictures twice). Next, we asked participants what they thought the purpose of the study was, and whether they used a specific method for matching names to faces. Finally, participants indicated whether they were familiar with any of the target faces and then provided their demographic details (age, gender, and nationality). To ensure that participants were matching names to unfamiliar faces, we discarded two data points (out of 2,420) because of participants' familiarity with the target faces. Including those points would artificially inflate the face-name matching rate.

Results and discussion. To analyze participants' accuracy in matching the true name to its face, we applied two methods. First, we created for each participant a proportion score that averaged his or her accurate choices and then computed the mean accuracy proportion for all participants. Using a one-sample *t* test analysis, we compared this mean accuracy proportion to a null hypothesis of a random-chance choice. Participants accurately matched the target's true name in 28.21% (SE = 0.93%) of the cases, which is significantly greater than chance (20%, $t(120) = 8.87$, $p < .001$, $d = .81$, 95% CI [.06, .10]). A similar pattern appeared in the second round; therefore, in all subsequent studies, we included only one round of judgments. Although straightforward and easily interpretable, these statistical tests do not account for variance due to the specific face stimuli used in the study (Judd, Westfall, & Kenny, 2012). Therefore, we conducted a second analysis using a mixed model that accounts for such variation, and ran a logistic mixed-effect regression model. Our model included the intercept as a fixed effect, as well as random intercepts for participant and item, which enabled us to examine whether the findings generalize across the sampled participants and stimuli. The results of the analysis showed the intercept to be estimated at $\beta = -.96$ (SE = 0.09) (corresponding to a raw value of 27.7% success rate). We then examined whether this value differs from chance level (20%, which corresponds to a value of −1.389 after logit transformation) by computing the standardized difference between the estimated intercept and chance level $\left(\text{i.e., } Z = \frac{\beta_{estimated} - \beta_{chance\ level}}{SE}\right)$. As predicted, the intercept differed significantly from chance level ($Z = 4.8$, $p < .001$; 95% CI [−1.13, −0.79], and in raw values [24.3%, 31.3%]). This finding shows that overall performance was indeed better than chance level, when considering both participants and items as random effects. Accuracy proportions for all faces in all studies appear in Table 1. Beyond the general support of a face-name matching effect, a variance is apparent by looking at the individual faces (see Table 1): For some targets, their names were chosen above chance level, whereas for other targets their names were chosen similar to or below chance level. This variance should be kept in mind, as it suggests the face-name fit as a possible characteristic of a person that should be studied. We return to this in the General Discussion.

Study 1A establishes the existence of the face-name matching effect by demonstrating a connection in real-life between people's face and name: For 17 of the 20 faces presented (85%), participants accurately matched the true name of the person depicted above chance level (20%), despite having no information other than the target face on which to base their choice.

Study 1B: Replicating Study 1A Using a Different Procedure for Filler Names

The purpose of Study 1B was twofold: We wanted to add a conceptual replication of Study 1A using a larger sample of stimuli, and refute an alternative explanation that the effect is driven by possible differences (such as differences in popularity) between target names and filler names. In this study, for filler names we used the pool of names of the other target faces in this study, instead of using an external pool of names that contained none of the target names, as in Study 1A. By using the other target names as filler names, we address the possible confounds related to the use of external filler names that suggest that target names have features that cause them to stand out (e.g., they are more/less popular than the filler names). Obviously, this procedure raises questions concerning the nonindependence of name choices and the repetitive appearance of the names, yet it served our goal of testing a possible confound of name uniqueness. In the other studies (Studies 2, 3, 5, 6, and 8), we continued to use the

Table 1
Binomial Test: Statistics for Each Target Face in All Human-Based Participant Studies

Target face	Observed proportion	Test proportion	*p* value
Study 1A			
Yael	.31	.20	.004
Alona	.31	.20	.003
Omer	.13	.20	.035[a]
Dani	.12	.20	.020[a]
Bat-Chen	.32	.20	.002
Amit	.26	.20	.052
Liran	.21	.20	.376
Dikla	.31	.20	.002
Ziv	.25	.20	.116
Aviad	.31	.20	.004
Alon	.38	.20	.000
Aya	.21	.20	.464
Sivan	.35	.20	.000
Haim	.36	.20	.000
Inbar	.31	.20	.004
Karin	.34	.20	.000
Dganit	.31	.20	.002
Amir	.34	.20	.000
Kfir	.36	.20	.000
Bar	.17	.20	.202[a]
Study 1B			
Aviv	.45	.25	.012
Avia	.18	.25	.247[a]
Ortal	.13	.25	.083[a]
Eyal	.19	.25	.312[a]
Ilan	.19	.25	.312[a]
Itamar	.32	.25	.229
Elad	.45	.25	.012
Ben	.23	.25	.473[a]
Gylad	.50	.25	.002
Gal	.25	.25	.568
Gilad	.32	.25	.229
David	.23	.25	.688
Dan	.22	.25	.432[a]
Dana	.06	.25	.999
Daniel	.46	.25	.008
Chen	.31	.25	.263
Tal	.42	.25	.021
Yehushua	.09	.25	.021[a]
Yonathan	.19	.25	.312[a]
Yael	.26	.25	.527
Yaakov	.16	.25	.176[a]
Liat	.36	.25	.128
Meir	.16	.25	.176[a]
Michael	.39	.25	.064
Maayan	.30	.25	.299
Matan	.32	.25	.229
Nadav	.39	.25	.048
Noam	.42	.25	.021
Nurit	.19	.25	.312[a]
Natalie	.40	.25	.051
Naama	.27	.25	.447
Netanel	.58	.25	.000
Sapir	.30	.25	.299
Adi	.19	.25	.312[a]
Omer	.49	.25	.003
Ofri	.07	.25	.008[a]
Omri	.21	.25	.394[a]
Anat	.30	.25	.326
Roe	.39	.25	.048
Rinat	.24	.25	.553[a]
Renana	.29	.25	.366
Rafael	.30	.25	.299
Sagi	.50	.25	.002
Shahar	.45	.25	.012
Shai	.19	.25	.278[a]
Shir	.23	.25	.473[a]
Shiran	.33	.25	.181
Shlomit	.21	.25	.394[a]
Shani	.30	.25	.299
Tom	.52	.25	.001
Study 2			
Claire	.35	.25	.008
Aurélie	.37	.25	.003
Alexis	.42	.25	.000
Clément	.35	.25	.008
Barbara	.39	.25	.001
Amandine	.48	.25	.000
Constance	.27	.25	.368
Véronique	.79	.25	.000
Pierre	.28	.25	.292
Julien	.34	.25	.014
Study 3			
Ilan			
Facial	.24	.20	.280
Control	.22	.20	.420
Ofir			
Facial	.24	.20	.274
Control	.25	.20	.197
Roe			
Facial	.49	.20	.000
Control	.33	.20	.015
Daniel			
Facial	.24	.20	.298
Control	.22	.20	.420
Omri			
Facial	.26	.20	.177
Control	.19	.20	.500[a]
Tal			
Facial	.24	.20	.298
Control	.16	.20	.316[a]
Dan			
Facial	.25	.20	.197
Control	.20	.20	.554
Maayan			
Facial	.11	.20	.065[a]
Control	.07	.20	.009[a]
Shir			
Facial	.27	.20	.121
Control	.20	.20	.554
Yael			
Facial	.16	.20	.316[a]
Control	.20	.20	.554
Study 5			
Israeli sample			
Ilan	.28	.25	.280
Ofir	.28	.25	.215
Roe	.42	.25	.000
Daniel	.41	.25	.000
Omri	.35	.25	.009
Tal	.45	.25	.000
Dan	.38	.25	.001
Maayan	.14	.25	.002[a]
Shir	.36	.25	.004
Yael	.21	.25	.189[a]
Claire	.24	.25	.421[a]

(*table continues*)

Table 1 (*continued*)

Target face	Observed proportion	Test proportion	*p* value
Aurélie	.22	.25	.249[a]
Alexis	.22	.25	.229[a]
Clément	.11	.25	.000[a]
Barbara	.24	.25	.421[a]
Amandine	.25	.25	.507[a]
Constance	.03	.25	.000[a]
Véronique	.21	.25	.214[a]
Pierre	.41	.25	.000
Julien	.30	.25	.132
French sample			
Claire	.33	.25	.035
Aurélie	.38	.25	.001
Alexis	.38	.25	.001
Clément	.27	.25	.366
Barbara	.38	.25	.002
Amandine	.37	.25	.004
Constance	.26	.25	.449
Véronique	.77	.25	.000
Pierre	.31	.25	.080
Julien	.35	.25	.013
Ilan	.34	.25	.024
Ofir	.17	.25	.032[a]
Roe	.18	.25	.047[a]
Daniel	.43	.25	.000
Omri	.17	.25	.023[a]
Tal	.08	.25	.000[a]
Dan	.33	.25	.034
Maayan	.50	.25	.000
Shir	.14	.25	.004[a]
Yael	.40	.25	.000
Study 6			
Ilan			
Facial appearance	.30	.25	.249
Hairstyle	.41	.25	.003
Inner facial	.31	.25	.175
Ofir			
Facial appearance	.39	.25	.009
Hairstyle	.27	.25	.380
Inner facial	.42	.25	.003
Roe			
Facial appearance	.48	.25	.000
Hairstyle	.53	.25	.000
Inner facial	.37	.25	.022
Daniel			
Facial appearance	.39	.25	.009
Hairstyle	.30	.25	.195
Inner facial	.17	.25	.083[a]
Omri			
Facial appearance	.39	.25	.009
Hairstyle	.31	.25	.175
Inner facial	.40	.25	.006
Tal			
Facial appearance	.34	.25	.064
Hairstyle	.29	.25	.280
Inner facial	.20	.25	.239[a]
Dan			
Facial appearance	.31	.25	.168
Hairstyle	.24	.25	.510[a]
Inner facial	.17	.25	.083[a]
Mayyan			
Facial appearance	.20	.25	.211[a]
Hairstyle	.26	.25	.490
Inner facial	.25	.25	.576
Shir			
Facial appearance	.39	.25	.009
Hairstyle	.33	.25	.081
Inner facial	.49	.25	.000
Yael			
Facial appearance	.44	.25	.001
Hairstyle	.33	.25	.081
Inner facial	.25	.25	.576
Study 8			
Ruven (name)	.53	.25	.009
Rubik (exclusive nickname)	.38	.25	.130
Gadi (name)	.38	.25	.130
Gadiel (exclusive nickname)	.07	.25	.101[a]
Hemda (name)	.05	.25	.031[a]
Himdul (exclusive nickname)	.00	.25	.002[a]
Oren (name)	.21	.25	.465[a]
Gomba (exclusive nickname)	.20	.25	.415[a]
Shmuel (name)	.48	.25	.021
Jacobs (exclusive nickname)	.11	.25	.111[a]
Chen (name)	.38	.25	.130
Chenesh (exclusive nickname)	.11	.25	.111[a]
Paul (name)	.06	.25	.050[a]
Polka (exclusive nickname)	.11	.25	.135[a]
Tal (name)	.22	.25	.519[a]
Tush (exclusive nickname)	.16	.25	.263[a]
Yulia (name)	.39	.25	.139
Bruri (exclusive nickname)	.11	.25	.135[a]
Irena (name)	.61	.25	.001
Iris (exclusive nickname)	.62	.25	.000
Adar (name)	.26	.25	.535
Lumier (exclusive nickname)	.14	.25	.192[a]
Ronit (name)	.48	.25	.021
Rona (exclusive nickname)	.16	.25	.263[a]
Ezra (name)	.10	.25	.091[a]
Moon (exclusive nickname)	.05	.25	.031[a]
Daniel (name)	.65	.25	.000
Noyman (exclusive nickname)	.37	.25	.175
Shira (No-nickname)	.50	.25	.001
Lotem (No-nickname)	.59	.25	.000
Ella (No-nickname)	.39	.25	.035
Elad (No-nickname)	.71	.25	.000
Ofir (No-nickname)	.38	.25	.054
Yonatan (No-nickname)	.58	.25	.000

[a] The observed proportion is smaller than the test proportion.

independent procedure used in Study 1A, ensuring an independent choice method.

Method.

Participants. Sixty-seven Israeli students from a local online panel (32 women, all native speakers, M_{age} = 25.59 years [*SD* = 3.36]) participated in the study in return for the equivalent of US$1.17 (three respondents gave incomplete responses and were therefore eliminated from the analysis).

Materials. We photographed 100 students who came to the lab in response to a general announcement. We used this pool of photographs in several of our studies while operating various screening methodologies, according to each study's purpose. In Study 1B, we filtered out targets who (a) had salient external features (e.g., a head covering that might indicate a religious background), (b) stated that they were not generally called by their given name, (c) were not born in Israel, and (d) had very uncommon names. This screening resulted in a set of 50 faces (23 female) that we used in this study. Similar to the targets used in Study 1A, all targets were young Israeli adults. In the purpose of easing participants' task and not overload them with 50 target faces, we randomly split the 50 faces into two sets of 25 target faces each, such that each target face appeared in one set, and each participant was presented with a total of 25 target faces. For filler names, we used the pool of names of the other target faces in the set. Different from Study 1A, participants were asked to match a name to its face by choosing one name among a set of four (not five) presented names (which also included the target's true name), thereby setting the chance level at 25%. For each face, we asked participants to select the true name of the person in the picture. Similar to Study 1A, we generated three different versions for each set, each containing the same filler names but for different target faces and in a different randomized order. That is, in each set, we randomly distributed the 25 true names among all faces, such that each face appeared with its given name along with three given names that belonged to other faces in the set. The set of possible target names was positioned below the headshot. For each participant, we randomized the order of the presented faces, as well as the location of the true name among the filler names.

Procedure. We introduced participants to the study as in Study 1A. We also informed participants that all the names belong to the targets and therefore they will encounter some names repeatedly over the course of the study. Before beginning the test trials that included the target faces, participants viewed the same single example and performed two practice trials. Participants then viewed 25 photographs of faces, one at a time, and were asked to decide which of the four names appearing below the photo was the true name of the person in the photo. We discarded eight data points (of 1,600) because of a participant's familiarity with a target face.

Results and discussion. This study replicated the findings of Study 1A, demonstrating a successful accurate match between names and facial appearances. Participants accurately chose the target's given name in 29.91% (SE = 1.11%) of the cases, which is significantly above chance (25%, $t(63) = 4.42$, $p < .001$, d = .55, with a 95% CI of the difference [.03, .07]). As in Study 1A, we conducted a logistic mixed effect regression model (with by-item and by-participant random intercepts) to confirm the findings of the t test analysis. Again, we found that overall performance significantly differed from chance ([25%; in logit space −1.098]; $\beta = -0.9$, corresponding to 29.0%, $SE = 0.09$; $Z = 2.32$, $p = .02$; 95% CI [−1.07, −0.72], and in raw values [25.6%, 32.6%]). Study 1B therefore replicates the face-name matching effect, with filler names of one target face being the true names of other target faces, thus addressing possible confounds related to the use of external filler names in Study 1A. This replication suggests that accurate choices represent a true face-name match and cannot be attributed to confounds related to filler names' features (e.g., they are more/less popular than the true names) that cause them to stand out.

Study 2: The Face-Name Match in a Different Culture

Study 2 aimed to replicate the face-name matching effect in a different culture to determine if the phenomenon is externally valid beyond one culture. In addition, this study aimed to address, differently than the method used in Study 1B, an explanation that the effect is driven by a possible difference between the targets' names and filler names, such as their popularity. Therefore, possible confounds related to ethnicity and name frequency were minimized.

Method. The study used new stimuli that were independently developed in France, holding ethnicity constant with a focus on Caucasian French targets and participants.

Participants. One hundred sixteen French participants (73 women, M_{age} = 23 years (SD = 2.90)) received €5 in return for their participation. The study was conducted in a university lab located in the center of a large French city.

Materials. We first selected 56 French given names that were neither too frequent nor too rare among the metropolitan population and had been similarly given since 1970 among Caucasian French people (INSEE 2011); we chose to go back more than 40 years to capture the social group in which our participants and targets primarily evolved. Using the lab subject pool and student directory of a large French university, we invited Caucasian students with one of these selected names to either send their picture or have their picture taken at the lab for the study. Using this procedure, we obtained a new set of 10 headshots (6 women). We used the other 46 names as fillers.

As in previous studies, headshots were presented one at a time. Similar to Study 1B, participants were asked to match a name to its face by choosing one name among four presented names (among which was the target's true name), thereby setting the chance level at 25%. For each face, we asked participants to select the true name of the person in the picture. As in Study 1A, the filler names did not belong to any face that appeared in the study, and all names—both the 10 true names and 46 filler names—appeared only once for each participant. We generated four different versions of the study, each presenting the filler names in a different randomized order, making sure each face never appeared with the same filler name among the different versions. Because we ensured no repetition of filler names per target face throughout the different versions, we used a total of 46 filler names. For each participant, we randomized the order of the presented faces, as well as the location of the true name among the filler names.

Procedure. We introduced participants to the study as in previous studies. Before beginning the test trials that included the target faces, participants viewed the same one example and performed two practice trials. No participant was familiar with any of the target faces.

Results. We replicated the face-name matching effect in a different country and culture: Participants accurately chose the target's given name in 40.52% (SE = 1.41%) of the cases, which is significantly greater than chance (25%, $t(115) = 11.00$, $p <$ .001, d = 1.02, 95% CI [.13, .18]). As in Studies 1A and 1B, we conducted a logistic mixed effect regression model (with by-item and by-participant random intercepts) to confirm the findings of the t test and examine whether overall performance was significantly better than chance level (25%, in logit space −1.098). Again, we found that overall performance significantly differed from chance ($\beta = -0.39$, corresponding to 40.3%, $SE = 0.2$; $Z = 3.56$, $p < .001$; 95% CI [−0.78, −0.003], and in raw values [31.4%, 49.9%]). Participants accurately chose the true name above chance level for all 10 faces (100%).

Study 3: The Face-Name Matching Effect Depends on Facial Information

The main goal of Study 3 was to confirm that the face-name matching effect stems from associating a name with facial information rather than from information conveyed only by the names (either true or filler). We wanted to control for any possible confound that is related to the characteristics of the names and which might lead people to choose a specific name, such as the length of the names, their frequency in the population, and so forth. By controlling for these confounds we were able to guarantee that participants' choice of the true name comes from its fit with the face, and nothing else. In addition, we aimed to control for stimuli age and ethnicity to rule out the possibility that these variables account for the face-name matching effect.

As in Study 1A, participants in Study 3 were asked to choose the person's true name among five suggested names for each headshot in a series of target faces. To ensure that the matching ability stemmed from facial information, and not from information related to the list of names presented below each face, we manipulated the facial information presented to participants. In particular, participants were either exposed to target faces as in Study 1A (facial condition), or they could not see the target faces because they appeared covered with a black square (control condition). Our prediction was that if it is facial information that causes the face-name matching effect, we should replicate the effect in the facial condition and fail to observe it in the control condition. Further, because we photographed targets, we firmly controlled for their age and ethnicity to rule out the possibility that these variables account for the face-name matching effect. Previous research shows that individuals associated a given name (Bruning, Polinko, & Buckingham, 1998; Lieberson & Bell, 1992) or facial feature (MacLin & Malpass, 2001) with certain socioeconomic components, such as ethnicity. For example, a specific ethnic group can be matched to a name that is more common in that group. Indeed, Chen, Gallagher, and Girod (2013) compared averaged faces computerized from photos and found ethnic and age relation between faces and specific names. For instance, an average computerized face of many women named "Alejandra" (often Hispanic) has darker skin and hair than the average face of women named "Heather" (often Caucasian), demonstrating an ethnic face-name fit. Both age and ethnicity play a role in our name and in our look. Our goal was to see whether the face-name matching effect could be demonstrated beyond these variables.

Method.

Participants. One hundred ten Israeli students from a local online panel (who did not participate in the previous studies, 54 women, all native speakers, $M_{\text{age}} = 26.56$ years [$SD = 3.28$]), participated in exchange for the equivalent of US$1.20.

Materials. We used the same pool of photographs used in Study 1B. From this pool of photographs, we identified targets (a) who belonged to a specific ethnic group (i.e., whose parents or grandparents were from Eastern Europe or Western countries; also termed "Ashkenazi") and (b) who were between 21 and 26 years old. This screening resulted in a set of 10 faces (four female) that we used in the actual study. We also ensured that the filler names were names for that age range and ethnic group. As in Study 1A, we generated five different versions of the study, each containing the same filler names but for different target faces and in a different randomized order. For each participant, we randomized the order of the presented faces, as well as the location of the true name among the filler names. The control condition only differed from the facial condition in that a black square covered the target faces.

Procedure. We randomly assigned participants to one of the two conditions. One example and two practice trials preceded the 10 target trials. As in Study 1A, we instructed participants to choose which of five proposed names was the true name of the person in each picture [behind the black square]. For an example of a trial choice, see Figure 2. To avoid participants in the control condition finding the

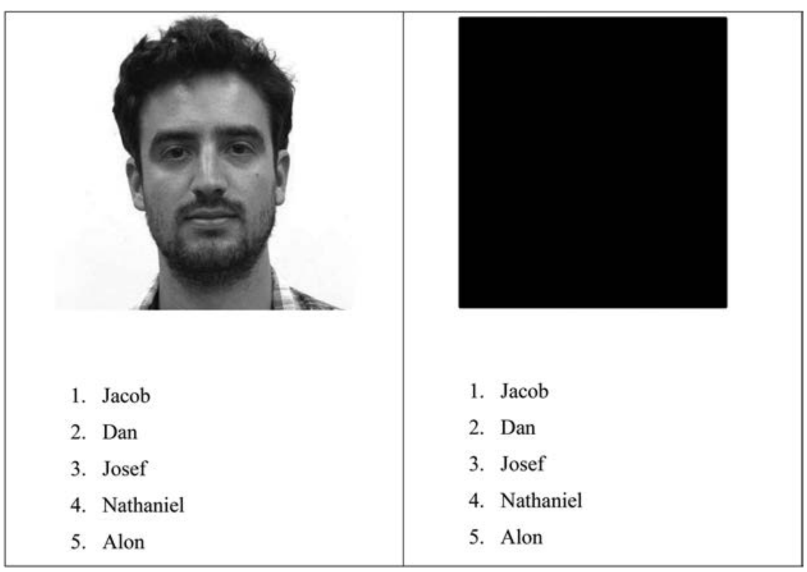

Figure 2. Example of a trial choice in Study 3. On the left is an example of a trial in the facial condition; on the right is an example of a trial in the control condition.

task odd, the black square was removed from each face after participants chose the target's name, and participants indicated whether the name they had chosen suited the face. Instructions were as follows:

> In each screen a photo of an individual (who is hidden behind a black square) will appear. Your task is to look at the list of names presented below the picture and determine the given name of the person (behind the black square). (After you choose a name, the black square will disappear, and you will see the person himself and state whether the name you chose indeed fits the face.)

No participant reported finding the task odd. After completion of the study, we detected a spelling error in one of the filler names for one of the target faces in one of the versions in the facial condition (one letter in one of the names was misplaced with a different letter). Including this stimulus might improve the results artificially, so we dropped it from the analysis. Including the stimulus indeed somewhat strengthens the effect (as detailed in the supplementary materials). To further ensure that target faces were unfamiliar and to avoid an artificial increase in the face-name effect, we discarded six data points (of 1,100) because in the debriefing questions participants indicated familiarity with the target faces.

Results and discussion. Figure 3 presents results from Study 3. Replicating our previous findings, participants in the facial condition accurately matched the target's name in 24.97% (SE = 1.93%) of the cases, which is significantly greater than chance (20%, $t(54) = 2.57$, $p = .013$, $d = .36$, 95% CI [.01, .09]). By contrast, as expected, in the control condition, participants failed to accurately match the target's name above chance level (20.44% (SE = 1.76%) vs. 20%, $p > .250$). We ran two logistic mixed effect regression models (with by-item and by-participants random intercepts) to compare overall performance to chance level in each of the experimental conditions (20%, or −1.389 in logit space). In the facial condition, overall performance was numerically better than chance, yet the estimated intercept failed to reach significance ($\beta = -1.14$, corresponding to 24.2%, $SE = 0.16$; $Z = 1.5$, $p = .133$; 95% CI [−1.46, −0.82], in raw values [18.8%, 30.6%]). As predicted, performance in the control condition was not different than chance ($\beta = -1.37$, corresponding to 20.2%, $SE = 0.13$; Z = 0.098, $p = .92$; 95% CI [−1.63, −1.12], in raw values [16.4%, 24.7%]).

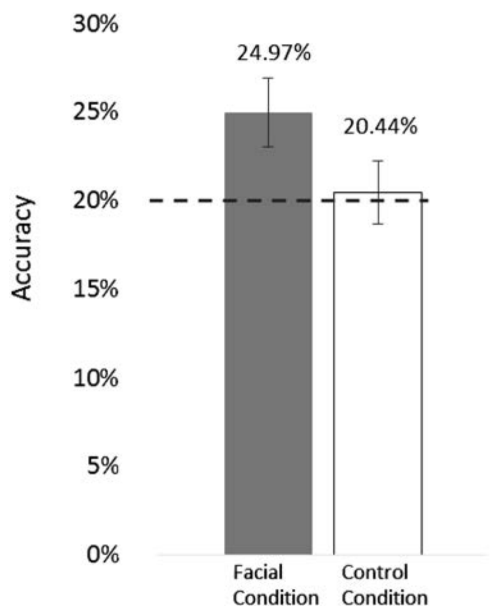

Figure 3. Results from Study 3: Mean accuracy ratings for matching the true name to its face, for the facial condition (where participants saw both the face and the name), and the control condition (where participants saw only the name). Error bars indicate *SEM*. The dashed line indicates chance level.

Following the above analysis, we cautiously propose that these different patterns across the two conditions suggest that it is the information conveyed in the target face that enables the face-name matching effect, and not any name-related confound. Participants accurately matched the true name above chance level for eight of the 10 faces (80%) in the facial condition, and only four of the 10 faces (40%) in the control condition. The probability of choosing the true given name was marginally significantly higher in the facial condition, where participants matched the true name to a face, compared with the control condition, where participants chose a name without being able to see the face, $t(108) = 1.73$, $p = .086$, $d = .33$, with a 95% CI of the difference [−.01, .10]. The fact that we found a face-name matching effect in the facial condition (using a t test analysis) and not in the control condition suggests that successful choices are not related to some feature of the names, such as the length of the names or their frequency in the population, which might lead people to choose a specific name. Importantly, these findings also indicate that our results cannot be attributed to information conveyed by the targets' socioeconomic background such as age or ethnicity.

Study 4: Even Computers Can Match—Eliminating Human-Biased Judgments

The main goal of this study was to examine whether nonhuman perceivers can also match names to faces, suggesting that the face-name match is independent of the measuring tool (in our case, human participants). For this purpose, we decided to use an entirely different paradigm—a computerized machine-learning approach. If a face-name matching effect emanates from a person's face, a computer should be able to learn it and make similar matches. This procedure also allowed us to significantly increase the number of stimuli, thus strengthening the validity of our behavioral studies.

In this computerized experiment, our goal was to show that the faces themselves convey enough information to produce the face-name matching effect, beyond any possible human bias. This is another way to rule out alternative explanations: When a human participant matches a name to a facial image, s/he uses personal, social, and historical contextual information associated with the names s/he is given as options. Yet, for a computer, a name is simply a category—its only uniqueness is that it is not any of the other names. In this sense, the computer is bias-free when it learns to match names from facial images; the only information it can access is the pixel values of the images it is presented with. Using a computer for matching the true name associated with a face enabled us to separate the information responsible for generating the phenomenon (i.e., facial appearance) from the tool we used for measurement (i.e., a human perceiver). Additionally, since the name has no meaning to the computer (it is just a label), various name characteristics (e.g., popularity) are irrelevant for the computer matching, thus excluding name related potential confounds. Finally, a computational approach allowed us to scale up the number of target faces, names, and comparisons used in the study, totaling 94,000 faces.

Experimental paradigm. We used an experimental paradigm taken from the field of *machine learning*, where a "learning" algorithm is trained to perform a matching task. Under this paradigm, an algorithm is presented with a set of examples (in our case, facial images) and their corresponding labels (in our case, the correct names). The algorithm then "trains" on this dataset and outputs a classification rule for matching the names of new, unseen facial images. The goal of this supervised training phase is to produce a classification rule that generalizes well; this means that it is trained with the objective of maximizing the accuracy of matching names of new unobserved faces. The accuracy of a classification rule is then measured on a held-out test dataset. In face-related computational tasks (such as face detection and recognition), learning algorithms often reach (and sometimes surpass) human-level accuracy. In our experiment, we used a class of learning algorithms called Deep Convolutional Neural Networks, which are the current state-of-the art for image-classification tasks (Taigman et al., 2014). More details on the learning procedure and algorithm can be found in the Appendix.

Because the goal of the current study is to replicate, extend, and expand the studies based on human participation, our experimental setting is designed to closely mimic the human experiments, under the constraints imposed by our computational approach. Similar to the human experiments, in each computerized trial an algorithm is presented with a facial image and two possible names (one of them being the true name of the person in the image and the second one being a false name, which is a name of a different target in the study). The algorithm's task is then to match the true one. We evaluated the performance of our algorithm by comparing its accuracy to chance level (0.5). Because a computer can easily handle a very large number of trials, it allowed us to thoroughly test all possible combinations of (same-gender) true and filler names with an exhaustive use of all available stimuli: The set of trials included all possible "true" names with all the possible "filler" names from the set of names we used, in both the training and testing phases. In this sense, "true" and "filler" lose their conventional meaning, since all names play both roles in different trials. Hence, for a pair of names presented to the computer to choose from, we refer to the "true" name as the "correct" name, and the "filler" as the "incorrect" name. Using only target names as filler names with no external names once again allows us to rule out some of the potential confounds due to differences between filler names and target name. As in the previous studies, we report per-name accuracies by averaging over all trials in which a focal name was presented as an option.

Data collection. To generate high-accuracy matches, a Neural Network (or any other learning algorithm) must be trained on a large set of training images. To meet this requirement, we collected a dataset of almost 100,000 profile images of members of a popular business-oriented social platform in France, all downloaded from the same public domain. We chose this as our source for facial images for two main reasons. First, compared with images from the general public domain, users of such social platforms tend to be relatively homogenous in terms of age, social status, and so forth. This allowed us to control for heterogeneity to some extent. Second, business-oriented profile images are typically "cleaner" in terms of pose, orientation, illumination, facial expression, and so forth. This is an important feature for successfully training an algorithm. Our dataset includes between 500 and 9,000 images for each of the 28 French names used in Study 2 as either a true or filler name. Table 2 details the number of images scraped for each name. Overall, our dataset consists of 36,013 facial images for 15 female names and 58,018 images for 13 male names.

Procedure. In each trial, the computer must choose one out of two names to match the presented face. To adhere to this task, we followed a standard training procedure known as "one vs. one," where for each (same-gender) pair of names, we trained a designated Neural Network. When two names were presented in a trial, the corresponding Neural Network was used for matching. To train a single name-pair Neural Network, we first randomly partitioned all images associated with those two names into a "training" set that included about 80% of the images, and a held-out "test" set with the remaining 20%. Then we trained our algorithm on the training set and evaluated its accuracy on the test set. We repeated

Table 2
Matching Accuracy of Names From Faces in the Computerized Study 4

Name	Accuracy	Chance	No. of images
Female			
Amandine	.59	.5	1,667
Angélique	.60	.5	941
Anne	.57	.5	5,738
Aurélie	.58	.5	2,733
Aurore	.59	.5	1,049
Barbara	.58	.5	1,053
Caroline	.57	.5	3,893
Céline	.57	.5	3,361
Charlotte	.58	.5	3,368
Claire	.58	.5	4,065
Emilie	.54	.5	2,246
Emma	.60	.5	830
Jeanne	.62	.5	583
Mathilde	.60	.5	2,389
Véronique	.65	.5	2,097
Average	.59		
Male			
Alexis	.57	.5	1,890
Arthur	.57	.5	1,140
Benjamin	.57	.5	3,014
Cédric	.58	.5	2,517
Clément	.60	.5	2,349
Jérôme	.61	.5	4,545
Julien	.58	.5	6,658
Laurent	.64	.5	5,764
Nicolas	.58	.5	8,905
Olivier	.63	.5	7,422
Pierre	.60	.5	8,148
Quentin	.60	.5	1,569
Romain	.59	.5	4,097
Average	.59		

Note. Displayed are the numbers of images scraped for each name. All reported percentages of accuracies are those from the held-out test faces, not from the training set of faces. All results are significant with $p < .001$ under a binomial test.

the above process for all pairs of names for several random train-test splits, while preserving the consistency of the samples across name pairs.

To reduce undesired visual variance, prior to training we used image-preprocessing techniques such as facial-area detection, facial alignment, resizing (to 50 × 50 pixels), and illumination normalization. We also converted the images to grayscale to further reduce counterfactual variance and algorithm complexity. All these are standard preprocessing steps in preparing pictures for training (e.g., Huang et al., 2007; Štruc et al., 2011; Viola & Jones, 2001).

Results and discussion. As in our human-participation studies, the trained learning algorithm was able to accurately match the true name significantly above chance level. Table 2 presents the face-name matching results for each true name, averaged over all trials in which that name was presented to the computer as one of the two optional names to choose from. In the following results, as well as in Tables 2–4, we note that all reported percentages of accuracies are those from the held-out test faces (and not from the training set of faces). All names were accurately matched to faces significantly above the 50% chance level with $p < .001$ under a binomial test. Matching accuracy ranged from 0.54% (Emilie) to 0.64% (Laurent). These results validate our theory that a match in the stimuli itself drives the face-name matching effects observed, and not any other possible behavioral bias of human participants.

Another advantage this experiment offers is the scalability of a computational approach, which allowed us to thoroughly test *all* possible combinations of (same-gender) true and filler names with an exhaustive use of all available stimuli. Tables 3 and 4 display the matching accuracies for all such pairs: All names were correctly matched with accuracy above chance for almost every other paired name. Specifically, Table 3 displays the accuracies of names matching for all female name pairs. Compared with the 50% chance level, the mean matching accuracy was 0.59%, with 0.43% as the lowest accuracy value (for Emilie as the correct name and Jeanne as the incorrect alternative), and 0.95% as the highest accuracy value (for Emma as the correct name and Véronique as the incorrect alternative). Table 4 displays the accuracies of names matching for all male name pairs. Compared with the 50% chance level, the mean matching accuracy was 0.60%, with 0.48% as the lowest accuracy value (for Alexis as the correct name and Arthur as the incorrect alternative), and 0.74% as the highest accuracy value (for Romain as the correct name and Laurent as the incorrect alternative, and for Laurent as the correct name and Romain as the incorrect alternative). These results suggest a wide-ranging phenomenon as opposed to one driven by only a small dominant set of easily distinguishable names.

Discussion of Section I—Testing the Face-Name Matching Effect

In Studies 1–3, we demonstrated the ability of social perceivers to accurately match faces and names, significantly above chance level, and beyond socioeconomic explanations. Specifically, in Studies 1A and 1B, using different methodologies for selection of filler names, we found that people can accurately identify a person's name based only on a headshot. In Study 2 we confirmed that the face-name matching effect can be replicated in the independent setting of a different country and culture. In Study 3 we aimed to rule out several alternative explanations and addressed methodological issues that might account for our results by controlling for possible confounds related to the names such as the presentation order, the method used to choose them, and their frequency in the general population. In addition, Study 3 used a tighter method to control for age and ethnicity of the target stimuli. The successful face-name matching rates suggest that the face-name matching effect occurs beyond methodological confounds, age, or ethnicity. Finally, in Study 4 we demonstrated that the information in the face itself is sufficient for a computer to generate the right match, thus eliminating human-biased judgments as an explanation.

To further minimize concerns regarding possible differences between the filler and target names leading to the face-name matching effect, we analyzed the control condition of Study 3, where participants see the list of names but they do not see the

Table 3
Accuracy of Name Matching in the Computerized Study 4 for All Female Name Pairs

True name	Filler name														
	Amandine	Angélique	Anne	Aurélie	Aurore	Barbara	Caroline	Céline	Charlotte	Claire	Emilie	Emma	Jeanne	Mathilde	Véronique
Amandine		.61	.67	.53	.56	.62	.62	.56	.53	.56	.57	.56	.64	.60	.66
Angélique	.55		.61	.56	.64	.60	.55	.62	.62	.57	.58	.58	.56	.61	.71
Anne	.62	.56		.56	.59	.54	.54	.58	.60	.57	.58	.59	.49	.61	.62
Aurélie	.57	.52	.58		.53	.55	.51	.58	.59	.59	.56	.60	.62	.62	.70
Aurore	.62	.58	.59	.53		.58	.50	.56	.54	.51	.61	.63	.61	.61	.72
Barbara	.60	.59	.56	.60	.59		.50	.60	.60	.52	.61	.57	.59	.62	.62
Caroline	.58	.59	.57	.61	.57	.57		.54	.57	.52	.57	.57	.60	.54	.61
Céline	.59	.51	.54	.53	.53	.63	.59		.63	.55	.58	.57	.45	.61	.65
Charlotte	.54	.59	.60	.57	.58	.63	.55	.60		.56	.56	.61	.56	.53	.70
Claire	.64	.58	.59	.57	.48	.54	.56	.58	.57		.59	.59	.55	.59	.63
Emilie	.58	.48	.60	.50	.50	.61	.51	.52	.52	.53		.58	.43	.57	.64
Emma	.59	.60	.63	.63	.58	.54	.55	.52	.61	.55	.52		.55	.60	.95
Jeanne	.63	.65	.68	.63	.58	.51	.59	.52	.59	.57	.57	.65		.87	.71
Mathilde	.58	.60	.65	.64	.56	.61	.63	.59	.57	.60	.58	.51	.59		.68
Véronique	.76	.66	.55	.67	.66	.54	.63	.63	.67	.62	.66	.74	.59	.73	

Note. Rows depict the true name in a trial; columns depict the filler name. All reported percentages of accuracies are those from the held-out test faces, not from the training set of faces.

Table 4
Accuracy of Name Matching in the Computerized Study 4 for All Male Name Pairs

	Filler												
True	Alexis	Arthur	Benjamin	Cédric	Clément	Jérôme	Julien	Laurent	Nicolas	Olivier	Pierre	Quentin	Romain
Alexis		.48	.60	.58	.55	.59	.51	.68	.55	.65	.59	.55	.54
Arthur	.55		.54	.58	.53	.67	.59	.69	.49	.60	.59	.54	.50
Benjamin	.54	.56		.58	.57	.60	.50	.67	.54	.65	.59	.54	.56
Cédric	.67	.56	.57		.66	.51	.53	.60	.51	.52	.64	.59	.61
Clément	.59	.51	.53	.65		.69	.56	.71	.60	.66	.61	.54	.55
Jérôme	.66	.66	.58	.60	.67		.60	.58	.58	.55	.56	.69	.64
Julien	.58	.56	.57	.57	.53	.59		.65	.56	.65	.61	.53	.57
Laurent	.61	.68	.68	.59	.72	.57	.67		.66	.55	.58	.69	.74
Nicolas	.54	.63	.55	.55	.56	.56	.54	.60		.59	.58	.60	.60
Olivier	.67	.59	.64	.57	.67	.57	.65	.57	.59		.61	.69	.69
Pierre	.60	.58	.59	.59	.61	.60	.61	.59	.58	.57		.61	.60
Quentin	.57	.51	.57	.67	.56	.62	.60	.69	.56	.73	.62		.53
Romain	.52	.58	.56	.60	.54	.64	.56	.74	.59	.67	.58	.53	

Note. Rows depict the true name in a trial; columns depict the filler name. All reported percentages of accuracies are those from the held-out test faces, not from the training set of faces.

targets' faces, which are hidden behind a black square. We examined the frequency of choice of each name in the control condition. Figure 4 demonstrates that the frequency of choice of the target names (in white) is distributed similarly to the frequency of choice of the filler names (in gray), thereby suggesting that the choice of name is not related to a name characteristic, but rather to information presented in the target's face. A *t* test analysis with the independent variable being the name type (target = 1, filler = 0) and the dependent variable the frequency of choice demonstrated that the mean frequency of both populations is similar (M_{target} = 11.2 vs. M_{fillers} = 10.9, $t(48) = .21$, $p = .835$), and thus name type cannot predict frequency of choice. This is of course a null test, which does not provide evidence, but we wanted to give a sense of whether the two distributions (of filler and target names) are indeed similar. As there is no difference in the frequency of choosing the filler and target names in the control condition, it seems that there is nothing about the names alone, both fillers and targets, that leads to the choice of the given name. It is only when participants see the face of the person that they are able to choose the true given name beyond chance level.

Together, these findings provide robust support for the existence of a face-name matching effect. They suggest the manifestation of a name in facial appearance: People look somewhat like their names.

Section II—The Role of Name Schemas in the Face-Name Match

The prior studies demonstrated the existence of the face-name matching effect using both human-based studies and computerized studies. Our next goal was to begin exploring the underlying process leading to this matching effect. Our hypothesis is that it is shared name schemas that produce a face-name matching effect in real-life. We suggest that name schemas are a necessary condition

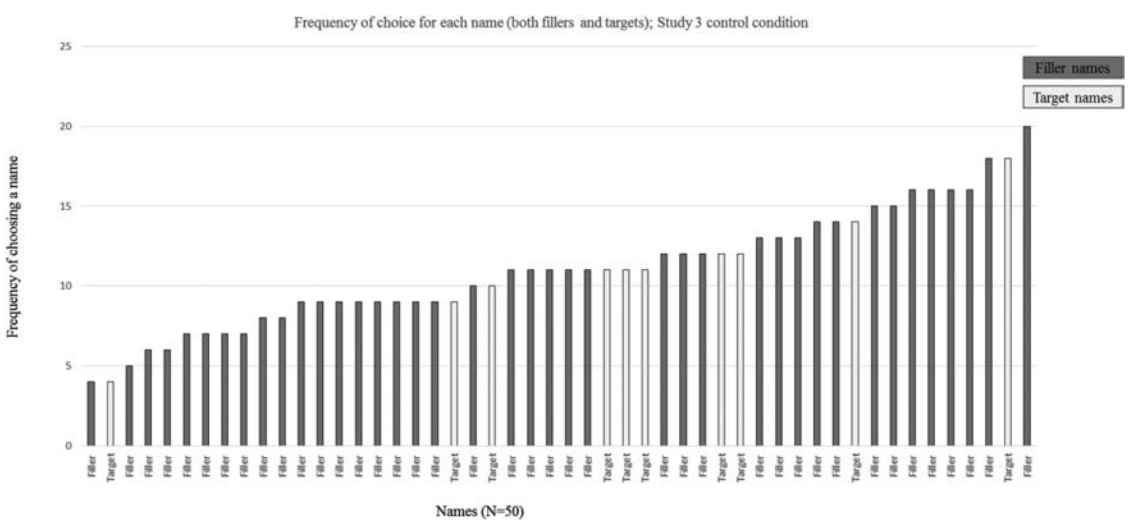

Figure 4. Results from Study 3 (control condition): mean frequency of choosing a name for filler names (in gray) and target names (in white).

both for the social actor's facial appearance to be influenced by his or her name as well as for the social perceiver to be able to detect a name in one's facial appearance. To test the role of shared name schemas we examined whether the name match exists when these schemas are absent from the perceiver's perspective. We expect the target person to look like his or her name, but the perceiver lacking this schema should lack the ability to match a name to its face.

To examine the lack of name schema for the social perceiver we conducted a cross-cultural study (Study 5) and expected an attenuation of the face-name matching effect when a name schema is from the other culture (i.e., the perceiver has no schema).

Study 5: A Cross-Cultural Study

The goal of Study 5 was to directly test the role of name schemas in real-life face-name matches, where the absence of a schema is from the perceiver's end. Our hypothesis is that in the case of the absence of a shared schema, social perceivers would not be able to accurately match a name to its face above chance level. For this purpose, we conducted a cross-cultural study, because people should be less familiar with names and/or faces from a different culture. Shared face-name prototypes, which are part of a name schema (i.e., the right facial appearance for a specific name), will exist within a culture in which people share familiarity with local names and faces through repeated exposure to them, thus developing a prototype of the "right" look of a person with a specific name. However, a perceiver from a different culture should not be familiar with these names and faces, hence will not have a face-name prototype. Because we suggest that the ability to accurately match a name to its face is based on shared name schemas, we hypothesize that being from the *same* culture as target faces/names should enable participants to accurately match a name to an unfamiliar face above chance level. However, being from a *different* culture than that of the target faces/names, for which schemas do not exist, should significantly attenuate participants' ability to have the "right" name in mind.

We tested this prediction by running participants in parallel in France and Israel. In both countries, participants matched French names to French faces, and Israeli names to Israeli faces. Our prediction was that we should observe the face-name matching effect when target faces and participants are from the same culture (replicating previous results); in contrast, the face-name matching effect should attenuate when targets are not from the participants' culture, thus strongly suggesting that the face-name matching effect depends on existing schemas.

Method.

Participants. In France, 112 French students (79 women, M_{age} = 22.75 years [SD = 2 .74]) participated in return for €4. Participants were not familiar with any of the faces in this study. In Israel, 123 Israeli students (62 women, M_{age} = 25.77 [SD = 4.82]) participated in return for the equivalent of US$2.30. We discarded two data points (of 2,460) because of participants' familiarity with a target face.

Materials and procedure. In both countries, participants matched the true name of 10 French targets (with French true and filler names) and 10 Israeli targets (with Israeli true and filler names). We predicted that we would observe the face-name matching effect when target faces are from the participants' culture (i.e., French [Israeli] participants should accurately match above chance level for French [Israeli] faces). By contrast, the face-name matching effect should attenuate when targets are from a different culture than that of the participants (i.e., French [Israeli] participants should be less accurate for Israeli [French] faces and names).

We used the same French and Israeli target faces and names as in Studies 2 and 3 (thus tightly controlling for age and ethnicity of the stimuli). The procedure was the same as in Study 2, except that participants matched the names of French and Israeli targets. The culture of the targets was a within-subject variable presented in separated blocks. We randomized the order of presentation of the French and Israeli blocks such that half the participants saw the Israeli faces first, whereas the other half saw the French faces first. Participants read that they would be presented with French [Israeli] people before the Israeli [French] set of faces. In both countries, we randomly assigned participants to the different study versions. For each target, four names appeared; therefore, the chance level in this study is 25%.

After participants matched all the targets' names, we asked them to rate their familiarity with each target name, using a 5-point scale (1 = *I do not know anybody with this name,* 5 = *I know many people with this name*). Because the main goal of Study 5 is to test people's ability to accurately match faces with given names of individuals of a different culture, we excluded observations of participants who stated that they knew a few or many individuals with a specific name from the other culture (i.e., they circled "4" or "5" on the familiarity scale). We thus excluded 50 observations (of 2,240) in the French study and 49 observations (of 2,460) in the Israeli study. Finally, we measured four emotions to control for any difference in affect while matching names from one's own culture versus another culture. Using 7-point scales anchored at 1 (*not at all*) and 7 (*very much so*), participants rated the extent to which they felt amused/content/annoyed and frustrated while matching the French names and while matching the Israeli names. We found that the French sample's emotions toward the Israeli stimuli were not related to accuracy (r = .133, p = .161) nor were the Israeli sample's emotions toward the French stimuli (r = .070, $p > .250$).

Results and discussion. We performed a 2 × 2 mixed-design ANOVA, with target faces' nationality (French vs. Israeli) as a within-subject factor and participants' nationality (French vs. Israeli) as a between-subjects factor. This analysis revealed no main effect for target faces' nationality, $F(1, 233) = 0.12$, $p > .250$, and a main effect for participants' nationality, $F(1, 233) = 10.85$, $p < .001$, as French participants correctly matched given names (M = 32.3%) significantly better than the Israeli participants did (M = 27.7%). As predicted, the face nationality x participant nationality interaction was found to be significant, $F(1, 233) = 79.61$, $p < .001$ (see Figure 5). To investigate this interaction, we examined the matching accuracy for each sample (French participants and Israeli participants) for each face nationality (French targets vs. Israeli targets).

French sample. As we expected, for French participants, the probability of accurately matching the target's name was significantly higher for French faces (37.95% (SE = 1.36%)) than for Israeli faces (26.68% (SE = 1.36%), $t(111) = 6.56$, $p < .001$, d = 1.25, 95% CI [.08, .15]). For French target faces, French partici-

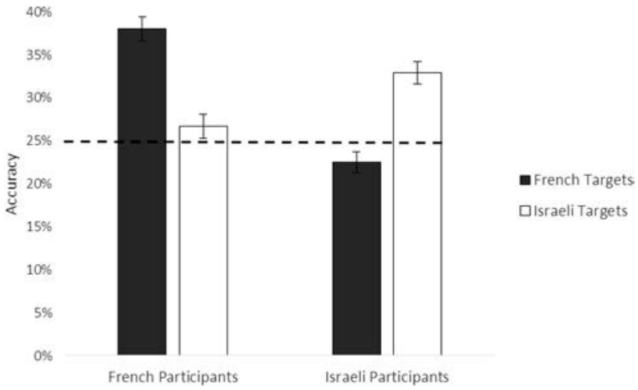

Figure 5. Results from Study 5: mean accuracy ratings for matching the true name to its face, for each sample (French vs. Israeli) and for each target's nationality (French vs. Israeli). Error bars indicate *SEM*. The dashed line indicates chance level.

pants accurately matched the true name significantly above chance level (37.95% vs. 25%, $t(111) = 9.55$, $p < .001$, $d = .90$, 95% CI [.10, .16]). By contrast, for Israeli faces, we no longer observed the face-name matching effect (26.68% vs. 25%, $t(111) = 1.24$, $p = .218$). A logistic mixed effect regression model (with by-item and by-participant random intercepts) confirmed the performance was significantly better than chance level (25%; in logit space 1.098) for French participants and French stimuli: $\beta = -0.5$, corresponding to 37.6%, $SE = 0.19$; $Z = 3.16$, $p = .002$; 95% CI [−0.87, −0.14], and in raw values [29.4%, 46.6%]. However, the difference from chance was no longer significant for French participants and Israeli stimuli: $\beta = -1.09$, corresponding to 25.1%, $SE = 0.24$; $Z = 0.02$, $p = .98$; 95% CI [−1.56, −0.63], in raw values [17.3%, 34.8%]. French participants accurately chose the true name above chance level for each of 10 French faces (100%) but for only five Israeli faces (50%).

Israeli sample. As expected, we observed the reverse pattern of results in Israel. For Israeli participants, the probability of accurately matching the targets' given name was significantly higher for Israeli target faces (32.90% ($SE = 1.31\%$)) than for French target faces (22.48% ($SE = 1.23\%$), $t(122) = 6.08$, $p < .001$, $d = 1.10$, 95% CI [.07, .14]). For Israeli faces, Israeli participants accurately matched given names significantly above chance level (32.90% vs. 25%, $t(122) = 6.01$, $p < .001$, $d = .54$, 95% CI [.05, .11]). By contrast, focusing on French target faces, Israeli participants matched the target's true given name in only 22.48% of the cases, which is significantly *below* chance level (25%, $t(122) = -2.05$, $p = .042$, $d = -.19$). A logistic mixed effect regression model (with by-item and by-participant random intercepts) yielded consistent results for Israeli participants and Israeli targets ($\beta = -0.74$, corresponding to 32.2%, $SE = 0.15$; $Z = 2.42$, $p = .016$; 95% CI [−1.03, −0.46], in raw values [26.3%, 38.7%]); as well as for Israeli participants and French targets ($\beta = -1.36$, corresponding to 20.4%, $SE = 0.22$; $Z = -1.2$, $p = .226$; 95% CI [−1.78, −0.94], in raw values [14.4%, 28.1%]). Participants accurately chose the true name above chance level for eight out of the 10 Israeli faces (80%) but only two of the French faces (20%).

Discussion of Section II—The Role of Name Schemas in the Face-Name Match

The above set of findings demonstrates the important role that shared name schemas have in the above-chance occurrence of a face-name real-life fit. Between cultures, we find an attenuation of the face-name matching abilities under conditions where shared name schemas are weak or lacking, suggesting that the social perceiver needs to be familiar with the schema.

Specifically, in Study 5, participants in two different cultures succeeded in accurately matching faces to names of individuals only if the target individuals depicted were from their own culture; they failed to match correctly when the faces and names were from a different culture, suggesting the critical role of familiarity with name schemas. Having a prior representation of the "right" face for that name is necessary. This carries two implications: First, we learn that successful face-name matching requires the existence of shared schemas. Second, these results are incongruent with a *bouba-kiki* figure-sound explanation, in which an innate face-name relationship is the only mechanism of the face-name match (e.g., a round face will be given a name that has a "round" universal sound, such as "Bob").

Section III—A Possible Underlying Mechanism: A Self-Fulfilling Prophecy

In Section III, we aimed to test a "self-fulfilling prophecy" as a possible account that could potentially explain the face-name matching effect. The idea of a self-fulfilling/defeating prophecy is not new, as it is already known that the way we are perceived, based on our facial information, affects how others treat us and in turn may influence our own self-perception and personality (Berry, 1991; Berry & Brownlow, 1989; Penton-Voak et al., 2006; Todorov et al., 2008; Zebrowitz, 1997; Zebrowitz & Collins, 1997; Zebrowitz, Collins, & Dutta, 1998). Nevertheless, our contribution is innovative in suggesting the possible effect on the way we look of a self-fulfilling prophecy driven by a social tag—one's name—which is believed to be chosen for us by our parents independently from our face. If this conjecture is valid, the association between faces and social perceptions could be a two-way street. According to this suggested process, people receive social signals related to their name, and they eventually come to look (to some extent) like their name as years pass. It has been found recently that the way a name sounds can be symbolically related to stereotypes associated with its social category (Slepian & Galinsky, 2016), and that a name can have drastic consequences for its possessor (Fryer & Levitt, 2004; Laham, Koval, & Alter, 2012). Consider the racial and gender associations of names. The exact same résumé receives more interviews when an applicant has a "White-sounding" name (e.g., Emily, Greg) compared with a "Black-sounding" name (e.g., Lakisha, Jamal; see Bertrand & Mullainathan, 2004). Similarly, academics perceive research applicants as more competent and hirable for a laboratory manager job if the applicant is named John versus Jennifer (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012). Moreover, based on given names, people can attribute to individuals characteristics with less clear boundaries, such as ethical caring, popularity, and successfulness (Mehrabian, 2001). Take, for example, an American girl named Katherine. Research found that people share a stereotype and perceive a woman with this name to be more successful than others (e.g., compared with a woman named Bonnie; Mehrabian, 2001). Thus, from the day she is born, a person named Katherine receives signals from society that relate to her name stereotype, leading her to act accordingly and eventually making her more and more similar to our shared representation of a stereotypical "Katherine." Thus, social signals related to a name might affect one's self-perception and personality, and to some extent, one's facial appearance.

If a self-fulfilling prophecy is the process responsible for the face-name matching effect, this means two essential things: First, there should be specific features in our face that people can control (even without being aware of it) to fulfill (or defeat) themselves according to the expectations society has for them vis-á-vis their name. Second, if for some reason one does not use her/his given name, there will be no expectations deriving from this given name and therefore no self-fulfilling prophecy, resulting in an attenuation of the face-name matching effect.

For the purpose of determining the potential role of prophecy, we first tested whether specific features of the face that are easily controlled are informative enough for the social perceiver (Study 6). One of the features of facial appearance that humans control most is their hairstyle, which is often used to represent and express one's identity. The hair is probably the largest facial appearance component and one that can be (and probably is) controlled, to some extent, by most individuals. Previous work has indicated that hairstyle is a particularly salient cue to distinguish between men and women (e.g., Roberts & Bruce, 1988), and more generally hair is important for facial recognition (Wright & Sladden, 2003). Therefore, in Study 6 we were interested to see whether peoples' hairstyle alone is enough to generate a face-name matching effect. Additionally, we wanted to explore whether a face-name matching effect can stem from the inner features of one's facial appearance. Though our inner facial features (e.g., face structure, eyes, and wrinkles) are less easily controlled compared with hairstyle, they are dynamic and active due to the numerous daily facial expressions one makes as a consequence of his or her personality or mood, or in response to social interactions. Indeed, Zajonc et al. (1987) demonstrated convergence in the inner facial features between spouses, and they proposed that emotional processes produce vascular changes that are, in part, regulated by facial musculature. The facial muscles are said to act as ligatures on veins and arteries, and they thereby are able to divert blood from, or direct blood to, the brain. An implication of the vascular theory is that habitual emotional use of facial musculature may permanently affect the physical features of the face (Waynbaum, 1907; Zajonc, 1985). The implication further holds that two people who live with each other for a long period of time, by virtue of repeated empathic mimicry, would grow physically similar in their facial features. Therefore, we were interested in exploring whether the face-name matching effect would occur even if we isolate the inner features of the face from the exterior features of the facial appearance, thus suggesting that the face-name fit also exists in the face itself.

In Study 7 we return to the machine-learning paradigm. Ideally, it would be informative to remove each and every facial component (similar to the hair removal in Study 6) and test which components are most important in individuals' ability to accurately match a name to a face. It is not possible to do such an examination using human participants, because faces without eyes or nose can look very weird and even creepy. However, a computer setup allows systematic removal of different parts of the face without causing an emotional bias measure when matching names to faces.

Finally, to test our hypothesis regarding the role of name usage in the ability to generate the face-name matching effect, we sought to test the case of people who typically do not use their given name but rather an exclusive nickname (Study 8). Using an exclusive nickname means a person is using his or her given name less. In such a case, the general name schema related to one's given name still exists for the social perceivers, yet the target person is not being called by this name and therefore the self-fulfilling prophecy process should not take place. The resulting lack of use of one's given name should translate into a weaker face-name match.

Study 6: A Possible Self-Fulfilling Prophecy—The Information Exists in a Controlled Component (Hairstyle)

Our primary goal in Study 6 was to test whether the face-name fit exists in easily controlled features of our facial appearance (in particular, hairstyle). A secondary goal was to explore whether the face-name match exists also in the face itself (excluding external features). To examine whether the matching ability stemmed not only from the full facial appearance but also from the hairstyle as well as from the face itself, we manipulated the facial information presented to participants. In particular, participants were exposed to target faces as in the previous studies (full facial appearance condition), or solely to the face of the target excluding any exterior feature (inner facial condition), or merely to the target's hairstyle (hairstyle condition). We predicted that all three conditions would manifest our initial social tag—our given name. Specifically, we expected that the face-name matching effect would occur when presenting a feature of the facial appearance that is heavily controlled by the individual—their hairstyle—as well as when presenting what may be perceived as less controlled but still alterable—our inner facial features; and of course when presenting the full facial information.

Method.

Participants. One hundred ninety-two Israeli students (94 women, all native speakers, M_{age} = 27.17 years [SD = 15.37]) from the same local online panel as previously (that did not participate in the previous studies) took part in this study in return for the equivalent of US$1.

Materials. We created three facial information conditions: full facial appearance, inner facial features, and hairstyle. For the full facial appearance condition, we used the set of 10 headshots (6 women) from Study 3, thus tightly controlling for age and ethnicity of the stimuli. We used the same filler names while randomly excluding 10 filler names, thereby presenting each target face with 4 names (the true name of the target and 3 filler names), thus setting the chance level at 25%. As in Studies 1A, 2, 3, and 5, the filler names did not belong to any face that appeared in the study, and all names—both the 10 true names and 30 filler names—appeared only once for each participant. We generated five new different versions of the study, each containing the same filler names but for different target faces and in a different randomized order. For each participant, we randomized the order of the presented faces, as well as the location of the true name among the filler names. The two additional conditions (inner facial condition and hairstyle condition) differed from the full facial appearance condition only in the portion of the facial appearance that was presented to participants. For this purpose, we used Photoshop software and eliminated different features of the facial appearance: In the hairstyle condition, we left only the external features of the facial appearance, which included mainly the hairstyle (together with one's ears and neck, as they are usually included in one's hairstyle), while the face itself was made indistinct. In the inner facial condition, we removed any feature of the facial appearance that is exterior to the inner part of the face, thus leaving an egg-shaped image of a face.

Procedure. We randomly assigned participants to one of the three conditions. One example and two practice trials preceded the 10 target trials. As in previous studies, headshots were presented one at a time, and participants were asked to match a true name to its face by choosing one name among four presented names (among which was the target's true name). As in previous studies, to ensure that target faces were unfamiliar and to avoid an artificial increase in the face-name matching effect, we discarded two data points (out of 1,920) because in the debriefing questions participants indicated familiarity with the target faces.

Results and discussion. A one-way between-subjects ANOVA was conducted to compare the effect of facial appearance on matching ability in the full facial appearance condition, hairstyle condition, and inner features condition. There was a marginally significant effect of facial appearance on matching ability for the three conditions, $F(2, 189) = 2.50$, $p = .085$. LSD post hoc comparisons indicated that the mean score for the full facial appearance condition (36.39%, SE = 1.94%) was significantly different from the inner features condition (30.19%, SE = 2.01%), $p = .027$. However, the hairstyle condition (32.78%, SE = 1.09%) did not significantly differ from the full facial appearance ($p = .194$) and the inner features conditions ($p > .250$).

We found that participants in the full facial appearance condition accurately matched the target's name in 36.39% (SE = 1.94%) of the cases, which is significantly greater than chance (25%, $t(60) = 5.87$, $p < .001$, $d = .75$, 95% CI [.08, .15]). Importantly, face-name matching effects were also found in the hairstyle condition (and the inner facial condition, using a t test analysis). In the hairstyle condition, participants accurately matched the target's name significantly above chance level (32.78% (SE = 1.90%) vs. 25%, $t(65) = 4.09$, $p < .001$, $d = .50$, 95% CI [.04, .12]). In the inner features condition, participants accurately matched the target's name significantly above chance level as well (30.19% (SE = 2.01%) vs. 25%, $t(64) = 2.58$, $p = .012$, $d = .32$, 95% CI [.01, .09]). A logistic mixed effect regression (comparing mean accuracy to 25% chance level; 1.098 in logit scale) yielded consistent results for the full facial appearance condition ($\beta = -0.56$, corresponding to 36.3%, $SE = 0.11$; $Z = 5.07$, $p < .001$; 95% CI [−0.77, −0.36], in raw values [31.6%, 41.2%]), as well as for the hairstyle condition ($\beta = -0.74$, corresponding to 32.3%, $SE = 0.12$; $Z = 3.01$, $p = .003$; 95% CI [−0.97, −0.50], in raw values [27.4%, 37.8%]), yet did not reach a significant effect in the facial condition ($\beta = -0.91$, corresponding to 28.7%, $SE = 0.18$; $Z = 1.08$, $p = .28$; 95% CI [−1.26, −0.56], in raw values [22.15%, 36.4%]). Achieving the successful matching accuracy in all three conditions (using a t test; and in only two conditions using a mixed model analysis [the full facial appearance and the hairstyle]) confirms that one easily controlled feature such as the hairstyle can be matched to one's name, and suggests that information conveyed solely by the target inner features might be enough to produce the face-name matching effect. The results of the matching of the inner facial condition led us to examine in the next study in more depth the role of targets inner facial features in the face-name matching effect. Participants accurately matched the true name above chance level for nine of 10 faces (90%) in the full facial appearance condition, nine of 10 (90%) in the hairstyle condition, and five of 10 (50%) in the inner facial condition.

Study 7: Key Facial Regions in Face-Name Matching

Given the results of Study 6, it is natural to ask what facial components play a significant role in the face-name associations. Theoretically, we could apply a procedure of removing facial components (e.g., eyes, nose, or mouth). However, such manipulations are likely to result in undesired effects on subjects' judgments. For instance, faces with no eyes tend to look weird, or even creepy. Another problem is that it is not clear what parts should be removed, or what the exact boundaries of a component are (e.g., where exactly the nose ends and the cheeks start).

To overcome these issues, we turned back to a computerized procedure to identify the facial regions that are key to the face-name matching process and convey most of the relevant information. Since an algorithm "perceives" an image simply as a collection of pixels, any loss in matching performance due to the removal of a facial component can be attributed directly to the corresponding missing pixels.

Method. In this study, we used the trained algorithms from Study 4 to generate *heat maps*—images in which each pixel's color represents its importance in matching. As described in Study 4, after the preprocessing phase, our images are facially aligned; this means that all significant facial parts (e.g., eyes, ears, nose) are roughly in the same location across all images. This feature allowed us to link a set of pixel locations (e.g., a square in the center of a facial image) with corresponding facial components (e.g., the nose) across all facial images.

Heat maps enabled us to examine what facial components are important (and unimportant) for correct matches across names. The basis for creating heat maps is measuring matching accuracy on obscured versions of the original images (Zeiler & Fergus, 2014). We take a facial image and "hide" a 7×7 pixel block centered at some pixel by coloring it gray. Our algorithm's task is (as before) to correctly match the name of a (now partially obscured) facial image. By "sliding" the obscuring block across all possible locations (that is, having the block centered at all possible pixels) and letting the algorithm generate a match for each such obscured image, we receive a per-pixel accuracy "map." This map is the same size as the original image, but instead of representing color, each pixel now represents whether the algorithm matched correctly (in which case the pixel value is 1) or not (in which case the value is 0). As such, 0's are considered "hot" (and colored red) because they reflect areas in the facial image that, if hidden, reduce the algorithm's accuracy. In other words, these "hot" areas are the ones that carry more weight in the face-name matching ability. In the same manner, 1's are considered "cold" (and colored blue) because they do not carry important information or classification. Hence, such maps are called heat maps.

We initially created a "base" heat map for every facial image used in a trial where the algorithm was successful in matching the true name. Then, we generated a final average heat map for each of the 28 names used in Study 4 by averaging all base heat maps from trials in which that name was the true name. To discard the asymmetry of pose in natural images, we reflect our resulting heat maps along the medial vertical axis (which results in left-right symmetric heat maps). As a result of averaging and the probabilistic output generated by the algorithm, the pixels in the final per-name heat maps contain values between 0 and 1, depicting the proportion of correct matches for when the pixel was the center of the obscured region. Analyzing these per-name heat maps allowed us to deduce the importance of different facial components in matching faces and names. A descriptive illustration of the heat-map generation process is presented in Figure 6.

Results. Figure 7 depicts the average heat maps for each name. To ease the visual association of heat values with facial regions, each heat map is superimposed with a grayscale image of the average face (the average is taken over the entire dataset). In most heat maps, "hot" red colors (which signify importance) were clustered and were typically concentrated at (one or more) specific facial parts, most of which were the parts that are related to our facial expressions (around the eyes and mouth). Consistent with a self-fulfilling prophecy account, this suggests that it is the more controlled features of our face that carry more weight in the face-name matching ability. Mildly hot areas (in yellow) were more spread out but were nonetheless concentrated in the face, whereas "cold" blue areas were mostly found at the margins. This validates our theory that the algorithm makes use of the information in the face, and not in its background or other nonfacial objects. Note that in contrast to Study 6, the heat maps did not show the hair as a significant component. This may sound surprising because in Study 6 hairstyle was found to be important, but in fact, the lack of significance in Study 7 is in line with the logic of our obscuring procedure. The reason for the (seemingly) unimportant role of hair in Study 7 is that the obscuring box is relatively small and hence never completely hides the hair from the algorithm. The computer is able to use ("see") most of the hair even when the obscuring box is placed on (part of) the hair, hence the hair is never removed. Overall, it can be seen from the heat maps that the significant information is almost always contained in a few distinct facial parts, which may vary across names. Future work could examine these differences.

Study 8: The Importance of Name Usage in the Face-Name Match

To test our hypothesis regarding the role of name usage in the face-name match, we sought to test the case of people who typically do not use their given name but an exclusive nickname instead. By exclusive nickname, we mean a nickname that is the only one of its kind (e.g., "Satchmo" was Louis Armstrong's exclusive nickname) and not a common nickname that is the short version of a name (e.g., "Bob" for Robert). Using an exclusive nickname means a person is using his or her given name less. In such a case, the general name schema related to one's given name still exists for the social perceiver, yet the person is not being called by this name and therefore the self-fulfilling prophecy process should not take place and s/he specifically should not look like his or her name. The resulting lack of use of one's given name should translate into a weaker face-name match.

Method.

Participants. We ran this study in Israel, where the use of exclusive nicknames is common. We ran the study in the lab and

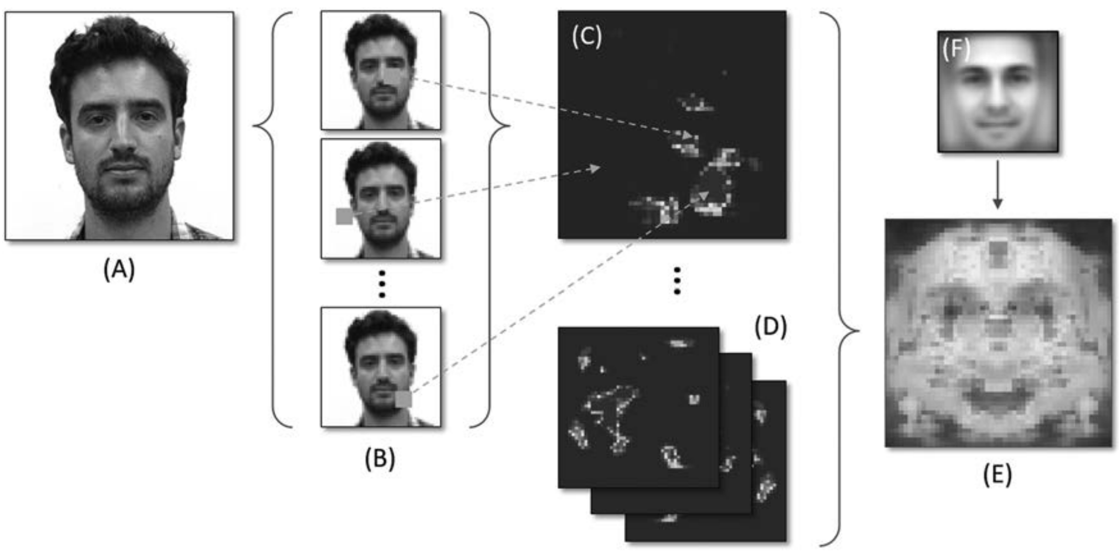

Figure 6. An illustration of the heat map generation process from Study 7. For each facial image (A), we create a set of obscured images (B), where an obscuring gray box is centered at all possible pixel locations. For each obscured image, the learning algorithm outputs a (probabilistic) match of the name. When the match is incorrect, the corresponding pixel in the heat map (C) of the current facial image is considered "hot" and is colored red, signifying importance for matching. When the match is correct, the pixel is "cold" and hence colored blue. This process is repeated for every facial image of the same name (D). Finally, all per-image heat maps of a certain name are averaged to produce the final per-name heat map (E), which is then overlaid with the average face (F).

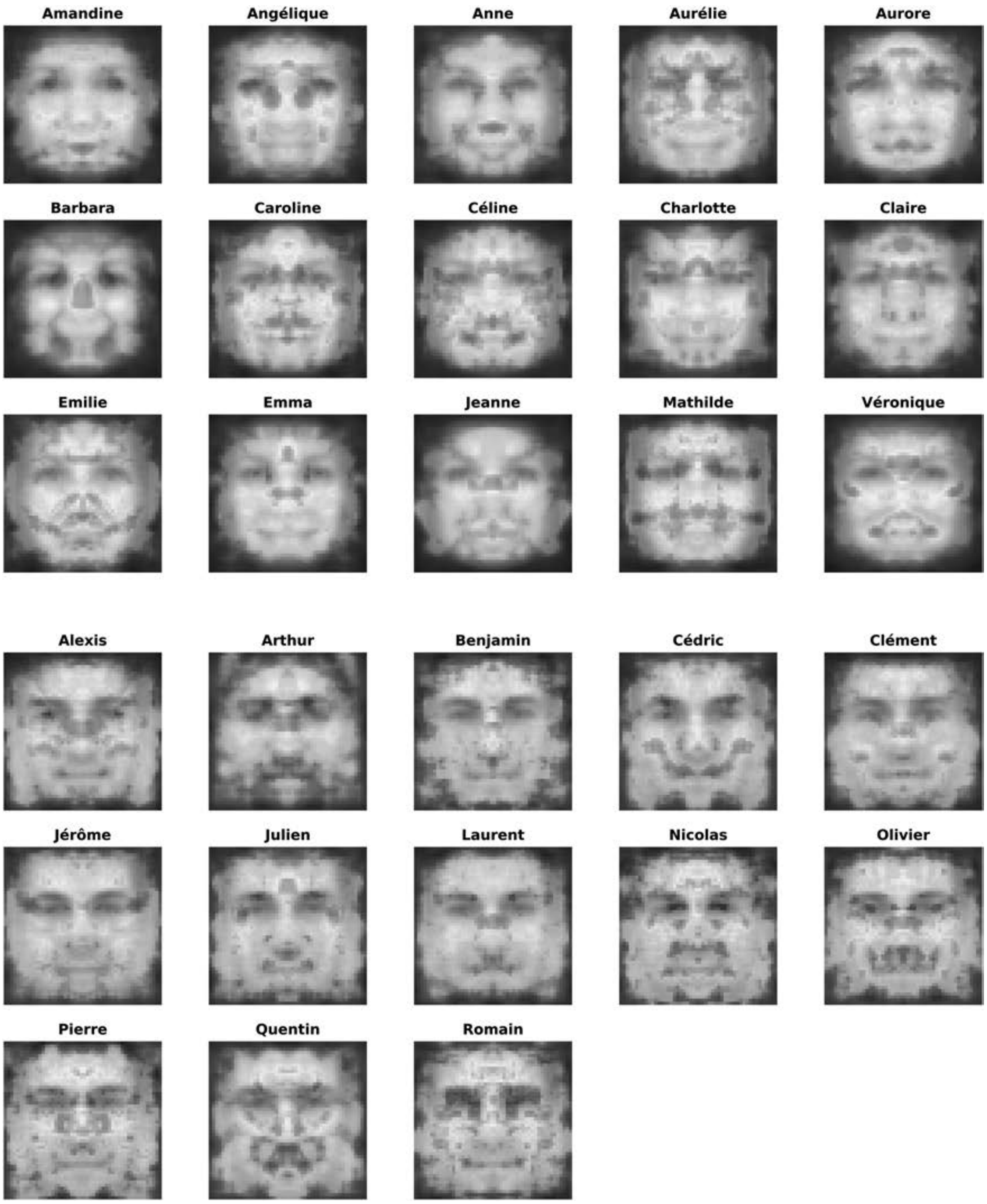

Figure 7. Heat maps for women (top 3 rows) and men (bottom 3 rows) from Study 7. Each heat map represents the areas in the images that are important (red or "hot") and unimportant (blue or "cold") for accurate matching. For ease of associating heat with facial regions, each heat map is superimposed with the average face.

collected 40 observations. Forty students (15 women, M_{age} = 24.68 [*SD* = 2.16]) from a large Israeli university took part in the study in return for the equivalent of US$2.60.

Materials and procedure. For target selection, we photographed students who responded to our announcement (posted on campus boards) inviting individuals with an exclusive nickname to participate. We photographed these individuals and had them answer a few questions regarding whether they used their nickname, whether they liked it, and additional demographic variables. We used 14 headshots of individuals who declared that most people refer to them by an exclusive nickname. Their exclusive nicknames were either exclusive distortions of their original given names (e.g., Rubik for Reuven), their last names (e.g., Noyman), or were totally unrelated names (e.g., Moon). None of the exclusive nicknames were related to the person's facial appearance. To this set of 14 faces, we added seven faces of targets who had no nickname (whether exclusive or not). We created six versions of the study. Whereas the seven no-nickname targets appeared in all versions with their true given name ("single given name"), the 14 targets with exclusive nicknames appeared either with their given name or with their exclusive nickname (counterbalanced between subjects). Thus, each participant viewed 21 faces (14 men, 7 women) in a random order that included seven faces with their single given name as the correct answer, seven faces with their given name as the correct answer, and seven faces with their exclusive nickname as the correct answer. For each target, four names appeared (two given names and two exclusive nicknames), only one of which applied to the target person whose face was presented; therefore, the chance level in this study is 25%. A hypothesis-blind research assistant randomly drew filler given names from known local websites containing catalogs of names, and made up exclusive filler nicknames. Note that importantly, the main comparison we were interested in was the choice of given names (and not exclusive nicknames) in the case of people who have exclusive nicknames compared with those who do not. As previously, one example and two practice trials preceded the test trials. We discarded 28 data points (of 840) because of participants' familiarity with the faces. Post hoc, we became aware that one of the seven single-name faces sometimes used a short version for his name; therefore, as a precaution we excluded this target face from our analyses, though doing so did not alter results.

Results and discussion. After averaging participants' accuracy for the three types of stimuli, a repeated measures analysis was conducted to compare the effect of name type on matching ability in the exclusive nickname condition, given name condition, and single given name condition. There was a significant effect of name type on matching ability for the three conditions, $F(2, 78) = 28.13$, $p < .001$. Repeated measures contrasts indicated that the mean score for choosing the true single given name was significantly higher (52.29% [$SE = 3.59\%$]), compared with matching names with faces of individuals who typically use an exclusive nickname and appeared with either their given name (34.91% [$SE = 3.30\%$], $p < .001$), or exclusive nickname (19.15% [$SE = 2.67\%$], $p < .001$).

Specifically, participants in the single given name condition accurately matched the target's name in 52.29% ($SE = 3.59\%$) of the cases, which is significantly greater than chance (25%, $t(39) = 7.6$, $p < .001$, $d = .1.20$, 95% CI [.20, .35]); Those in the given name condition matched the target's name in 34.91% ($SE = 3.30\%$) of the cases, which is significantly greater than chance (25%, $t(39) = 3.01$, $p = .005$, $d = .48$, 95% CI [.03, .17]); Finally, as expected, in the exclusive nickname condition, participants failed to accurately match the target's name (19.15% ($SE = 2.67\%$), which is significantly below chance level (25%, $t(39) = -2.19$, $p = .035$). Results were similar using a logistic mixed model regression: An above-chance (25%; 1.098 in logit space) performance was obtained in the single name condition ($\beta = 0.10$, corresponding to 52.6%, $SE = 0.21$; $Z = 5.6$, $p < .001$; 95% CI [−0.31, −0.53], in raw values [42.1%, 62.8%]). We found no effect in the given name condition ($\beta = -0.81$, corresponding to 30.8%, $SE = 0.3$; Z = 0.95, $p = .34$; 95% CI [−1.4, −0.22], in raw values [19.7%, 44.6%]), and a significant below chance-level performance in the nickname condition ($\beta = -1.81$, corresponding to 14.1%, $SE = 0.35$; $Z = -2.04$, $p = .04$; 95% CI [−2.49, −1.13], in raw values [7.7%, 24.5%]). Participants accurately chose the true name above chance level for all faces (100%) in the case of target faces with no nickname; nine of 14 faces (64%) in the case of targets with exclusive nicknames who appeared with their given names; and three of 14 faces (21%) in the case of targets with exclusive nicknames who appeared with their exclusive nickname (accuracy proportions for each name or nickname appear in Table 1). Note that including the target face we excluded above in our analyses did not change our results.

The significant increase in the accurate matching ability for individuals who use their given name more frequently compared with those who use their name less frequently suggests the possibility that the mechanism responsible for the face-name matching effect concerns the influence of one's name on facial appearance. There is also a possibility that a person received an exclusive nickname because his or her given name did not match him/her to begin with. We return to this conjecture in the General Discussion. Overall, the relatively high levels of match of the given names may also be a result of a low tendency to choose an exclusive nickname. Though participants chose exclusive nicknames to be the true names in some cases, the frequency of such choices was lower than that of given names. This tendency may reflect a difficulty in associating a name with its face when the name is unique. Such difficulty may result from one's tendency to avoid choosing a "unique" answer, or it could reveal a deeper pattern whereby the social perceiver is not able to generate associations when s/he lacks a schema for them. The fact that participants typically did not choose (rare) exclusive nicknames suggests that in reality the chance level in this study is above 25%, and it might be closer to 50% because each face was presented with only two given (non-nickname) names. Notably, it is critical to emphasize the key finding in this experiment: that, in spite of this issue, the difference between the single-name condition (e.g., faces of people who do not have a nickname) and the given name condition (e.g., faces of people who typically use nicknames) is significant.

Discussion of Section III—A Possible Self-Fulfilling Prophecy

In Section III, we aimed to further understand the process responsible for the face-name matching effect. Results from Studies 6, 7, and 8 provide initial evidence that the mechanism we seek concerns the influence of one's name on one's facial appearance. That is, one's name might serve as a self-fulfilling prophecy that is eventually realized in one's facial appearance.

First, we demonstrated that one of the features of facial appearance that humans control most—their hairstyle—is informative enough for social perceivers to be able to accurately match a hairstyle and a name (Study 6). Additionally, we suggest that the face-name matching effect also stems from within the face itself, excluding exterior features such as hairstyle, ears, neck, and so forth. We used heat maps to further explore the importance of different facial regions, analyzing which facial regions are important for accurate matches of the learning algorithm used in Study 4 (see Study 7). The heat maps showed that the significant information is contained in a few distinct facial parts, usually the ones that are responsible for our facial expressions. Finally, our data point to a possible self-fulfilling prophecy explanation behind the development of face-name matches that end up producing face-name matching effects, since the importance of name usage was demonstrated (Study 8).

An Aggregated Mixed Model Analysis With Faces as a Random Factor

Given the large number of studies in this research, we enjoy a large number of stimuli across all studies. We used more than 100 different target faces in the human-subject studies, and 94,000 faces in the computerized studies. We ran an additional aggregated mixed-effect analysis that accounts for variations due to stimuli in addition to interindividual variations, treating both stimuli and participants as random factors (Judd, Westfall, & Kenny, 2012). To do so, we combined data from all the human-subject studies in which we predicted the face-name matching effect. To be conservative, we kept the computer study outside this analysis, as the extremely large number of stimuli in this study would bias the results in our favor. For the same reason we also did not include Study 8, as it seems that participants were not inclined to choose unique names, thus altering the chance level (including Study 8 in this analysis yielded similar significant results). Our analysis therefore includes Studies 1A, 1B, 2, 3 (facial condition), 5 (from this cross-cultural study we included only the two conditions where we expected to see a face-name matching effect: French participants who viewed French targets and Israeli participants who viewed Israeli targets), and 6.

We ran a logistic mixed-effect regression model, with fixed effects for the intercept and for the study (effect-coded, to maintain the interpretation of the intercept as overall performance), as well as random intercepts for subject and stimuli. The results of the analysis showed the overall intercept to be estimated at $\beta = -.75$ (corresponding to a raw value of 32.1% success rate; $SE = 0.08$). We next compared this value to the overall chance level, across the full aggregated dataset. Because different experiments included in the analysis had different chance-level thresholds, we calculated an aggregated chance level, defined as the mean chance level weighted according to the number of trials in each experiment. The aggregated chance level was found to be set to 23.5% (corresponding to a value of -1.18 in logit space). To examine whether performance was better than this chance threshold, we next calculated the standardized difference between the estimated intercept and chance $\left(\text{i.e., } \frac{\beta_{estimated} - \beta_{chance\ level}}{SE}\right)$. As predicted, the intercept differed significantly from chance level ($Z = 5.11$, $p < .001$; 95% CI [-0.92, -0.59], and in raw values [28.6%, 35.7%]). This means that overall performance across experiments was significantly better than chance level. In sum, in addition to the results we present for each of our human-subject studies, an aggregated analysis including a large number of stimuli while accounting for the variance in those stimuli across studies provides strong support for the existence of the face-name matching effect.

General Discussion

In eight studies using color headshot photographs, each appearing with the target's given name (or nickname, Study 8) and filler names, participants and a computer were able to accurately identify above chance level the true given name of people. We eliminated possible alternative explanations and addressed methodological issues: First, we demonstrated this face-name matching effect using external names as filler names (Studies 1A, 2, 3, 5, 6 and 8) as well as other targets' names as filler names (Studies 1B and 4). Study 1B ruled out the possibility that the frequency of the names (or other feature related to the targets' names) drives the observed effect, because participants correctly matched the names to their corresponding faces even when the same names appeared with different faces. We replicated the face-name matching effect in a different country and culture (Study 2). When face information is absent (Study 3, control condition), the face-name matching effect disappeared (participants were unable to choose the true name above chance level). These results suggest that accurate face-name identification in the facial condition (with a *t* test analysis, but not with the mixed model analysis), as well as in the other studies (with both types of analyses), is attributable to information conveyed by the face and not to possible confounds related to the names, such as the way we chose the name, the way the names were presented, the frequency of the listed names, and so forth. The results of the control condition in Study 3 are of course a null effect, which does not constitute evidence by itself, but they do give us a sense of the nature of the face-name matching effect's dependency on facial information rather than on choices that depend solely on names. Additionally, because there was no difference in the frequency of choosing the filler and target names in the control condition of Study 3, it seems that there is nothing about either the target or filler names per se that leads to the choice of the given name. It is only when participants see the face of the person that they are able to choose the true given name beyond chance level. Furthermore, participants accurately identified the correct name when tightly controlling for age and ethnicity of the stimuli (Studies 3, 5, and 6), implying that successful identification goes beyond socioeconomic cues such as age or ethnicity. Finally, using a computerized experimental paradigm (Study 4), we completely removed the human factor from the matching process, strengthening the claim that the facial images themselves contain the relevant information. A computerized approach also allowed us to use many target images, thus strengthening the generality of the phenomenon. Together, it is noteworthy that our studies entail a multimethod approach and thus comprehensively address the possible confounds having to do with filler names differing from target names. Possible differences between the filler names and target names constituted a key concern regarding the face-name matching effect. Accordingly, Studies 1B, 3, and 4 were designed to directly test this possibility.

The Important Role of Face-Name Prototypes in Generating the Face-Name Matching Effect

Together, the first four studies suggest that we look like our name. What can explain the face-name matching effect in real-life? The first main factor that we suggest—for the social actor to form the face-name fit as well as for social perceivers to be able to detect it—is the existence of shared face-name prototypes. Study 5 offered initial evidence for this critical role of shared face-name prototypes, as the face-name match was found to be culture-dependent: Participants were unable to match names to faces from another culture. Having shared face-name prototypes requires one to be familiar with the facial appearance of people with a specific name. This evidence offers an indication of the critical role of shared face-name prototypes in the manifestation of actual face-name matches and the ability to accurately match a name to its face.

A Possible Self-Fulfilling Prophecy

A plausible process explaining the face-name matching effect is that a self-fulfilling prophecy is responsible for the development of actual face-name matches. The process may be of a direct nature such as the expectation of a specific look for a specific name and/or occur via the mediation effect of a name on personality. One way or the other, our facial features may change over the years to eventually represent the expectations of how we should look.

First, we find that one of the facial features that humans control most (their hairstyle) is informative enough for social perceivers that individuals were able to accurately match hairstyle and a name (Study 6; note that when both facial information and hair information were available, the matching performances were at the maximum level). Second, the fine-grained heat-map analysis (Study 7) of the learning algorithm shows that each name has specific facial parts that are crucial for producing a correct match, most of which are the parts related to our facial expressions (e.g., around the eyes and the mouth). Finally, comparing people who use their name more with people who use it less (i.e., use an exclusive nickname), Study 8 demonstrates the primacy of name usage in the development of face-name matches. If the infrequent usage of one's given name significantly decreases the face-name matching effect (i.e., as it does in the case of using an exclusive nickname), this implies that usage is important for manifesting a face-name match, thus strengthening the argument for self-fulfilling prophecy as the leading mechanism of the face-name matching effect.

Notably, for social perceivers to be able to accurately match names to faces, two things must occur: They have to be familiar with the face-names prototypes, and the target's facial appearance must reflect this prototype. Looking at results from Study 8, we see that if the correct face-name matches were solely dependent on the fit between the target's face and name, then using an exclusive nickname would have led to accurate identifications of the *exclusive nickname* of the target. However, this was not the case because more is required than the fit between a face and a name; in addition, the social perceivers (i.e., the participants) must share a common perception of relevant social schema, and they would not be familiar with the exclusive nicknames, as these are usually unique and rare. On the other hand, if the only thing that drives the effect is the existence of face-name prototypes, then the social perceivers (i.e., the participants) should have been able to accurately identify the *given name* of the target, even if he or she does not use it. Nevertheless, perceivers performed significantly worse in identifying these cases compared with the cases of targets with no nickname, which can lead us to conclude that although schemas exist for the given names of the targets who have an exclusive nickname, if they do not use the given name, the perceivers are not able to identify it, and this points toward the process of a self-fulfilling prophecy.

Then again, it could be that people who have exclusive nicknames received them because their given name did not fit them well enough to begin with, and this is the reason for the face-name attenuation in the case of exclusive nicknames. This could suggest a different explanation of the face-name match. Nevertheless, because names are often chosen before birth, or at least before the person looks much like he or she will as an adult, even if there could be some influence of the facial features of a baby on the name she or he is given, we believe this process is not likely to be the central mechanism behind the face-name matching effect. Moreover, we wish to emphasize previous findings which support the notion that in the first period after birth, babies are homogeneous in appearance compared with adult facial appearance and not much may be inferred from their photographed faces alone (Gordon & Tanaka, 2011; Yovel et al., 2012; we note interesting evolutionary perspectives regarding this topic in Alvergne, Faurie, & Raymond [2007] and McLain et al., [2000]). Even gender recognition at these stages is relatively difficult (see, e.g., Kaminski et al. [2011] and Tskhay & Rule [2016]). Gender, which is a basic differentiator among humans, is recognized better for adults than for newborns, suggesting that if any social inference may be derived from baby faces, it is presumably weak. Still, future work should examine the face-name matching effect for babies (for instance, some people see a baby and then decide to replace the original name they planned to give him/her, for a better fit).

Additionally, results of the cross-cultural study (Study 5) negate, to some extent, a *bouba-kiki* face-name inborn explanation as the stand-alone mechanism: Considering the possibility of an inborn process, one link can be a universal match between one's look at birth and the sound of the name he or she receives (similar to the *bouba-kiki* effect where the rounder shape is universally called by the rounder sound "bouba"). However, if this were the case, then a universal identification would be possible not only within a culture but also between cultures, and our findings demonstrate the inability to match names to faces from a different culture (Study 5). If people from another culture are not able to accurately match faces to names, this means the effect is not based on a universal sound-shape relationship.

Taken together, our findings point to the existence of a face-name matching effect in real life and to the possibility that the expectations related to name schemas eventually are realized in the outer world by a type of Dorian Gray effect, such that people will develop over time a look that somewhat resembles associations common to their name. Future work should continue examining the precise nature of the mechanism leading to the emergence of this face-name matching effect. One developmental test could involve examining the ability to match a target's name to his or her face in different stages of his or her life, or comparing the ability to match faces and names for targets who are babies with that of targets who are adults. Such investigations would involve new methodologies (e.g., using both stimuli as well as participants from different age groups).

Contribution to Social Structuring

The general idea of a self-fulfilling prophecy is not new, as it is already known that the way we are perceived, based on our facial information, affects how others treat us and in turn may influence our own self-perception and personality (Berry, 1991; Berry & Brownlow, 1989; Penton-Voak et al., 2006; Todorov et al., 2008; Zebrowitz, 1997; Zebrowitz & Collins, 1997; Zebrowitz, Collins, & Dutta, 1998). Nevertheless, our contribution is innovative in demonstrating the possible effect on the way we look of a self-fulfilling prophecy driven by a social tag—one's name—which is believed to be chosen for us by our parents independently from our face. If this conjecture is valid, the association between faces and social perceptions could be a two-way street. Although earlier research focused on the effect that one's look has on social perceptions, we focus on the effect of social perception on one's look and demonstrate that actual face-name matches are possible. That a name can be manifested in appearance, even to a small extent, emphasizes the important role of social structuring in interaction between the self and society. No doubt there remain additional explanations or factors responsible for one's name being manifested in one's facial appearance that future research should address.

In an interesting recent article, Kramer and Jones (2015) reported results that may seem to contradict ours. Specifically, in one study they found evidence that American women were able to identify American faces from names at significantly above-chance levels of accuracy. However, in a second study they failed to replicate this result with a new set of British faces and participants, leading the authors to conclude that they found no overall evidence suggesting an association between names and faces. Notably, they used a different methodology, presenting participants with *two* faces that appeared with *one* name and asking, "Which of these individuals is named X?" We believe that this different methodology might elicit a different processing of the faces. One possibility is that comparing two faces (rather than names as in our paradigm) leads to a feature type of processing rather than a holistic one. In turn, this may hinder impression and judgments from the faces (Meissner & Brigham, 2001; Peterson & Rhodes, 2003; Tanaka & Farah, 1993). Also, it could be that comparing only two independent options (rather than four or five as in our paradigm) leads to a different type of comparison and processing (Ganzach & Schul, 1995). Future studies could compare the two different methodologies using the same stimuli.

Crucially, our findings demonstrate a tangible effect of stereotypes in real life. Not only are names perceived as stereotypical, but these stereotypes are activated and potentially affect one's facial appearance. The fact that perceivers can choose the correct name just by looking at a face demonstrates that these stereotypes can impact one's very being—a powerful example of social structuring. Social structuring literature has demonstrated that people apply cultural stereotypes to their self (Crocker, Major, & Steele, 1998; Greenwald et al., 2002; Swim & Stangor, 1998; Wheeler & Petty, 2001). Nevertheless, these activated stereotypes were usually limited in two ways: First, researchers have shown that activating stereotypes can influence people's behavior (not their look); and second, they focused mainly on the *social* identities of gender, age, and ethnicity (and not their characteristics) (for a review see Wheeler & Petty, 2001). Uniquely, the face-name match is a demonstration of a stereotype influencing our physical identity, and this influence is beyond the studied stereotypes of gender, age, and ethnicity in that a name is something arbitrary that is chosen for us at birth. Considering these special characteristics of faces and names, we believe the face-name matching effect demonstrates a most powerful social structuring that has thus far been ignored.

Contribution to Social Psychology

The existence of actual face-name matches may have cognitive and social psychological implications. Looking at the matching ability across the different faces in the various studies (see Table 1), one sees variation among the faces: For some targets, their faces have a great fit with their name, whereas for other targets their face does not fit their name at all. Clearly, some people have a name that is easily matched to their face, while others do not. These varying levels of congruency that people have between their face and name may represent another individual characteristic. Future work could explore the origins and consequences of this possible face-name congruency characteristic. The various levels of face-name fit could shape, for example, face and name memory (Pantelis, van Vugt, Sekuler, Wilson, & Kahana, 2008). Interestingly, in a learning face-name task, in one condition, the experimenter arbitrarily chose names for faces, whereas in the other, participants chose names that "fit" the faces. Participants recalled better the names that they assigned compared with names arbitrarily assigned (Cohen & Burke, 1993). Along these lines, when morphed faces and names were not ideally matched, participants learned the matching more slowly than when faces and names were well matched (Lea et al., 2007). These differences may be the result of the congruency level of the face and name, according to the existing shared face-name schemas, affecting one's encoding and then recall of the information. Therefore, the different level of fit between a person's name and his or her facial appearance may have important consequences. For example, a good face-name fit may be more fluent (Reber, Schwarz, & Winkielman, 2004) and therefore influence perceptions of attractiveness, trustworthiness, and even compatibility of job candidates.

Conclusion

The current set of studies suggests that the association between faces and social perceptions could be a two-way street. Whereas earlier research focused on the effect of facial appearance on social perception, we focus on the effect of social perception on facial appearance and demonstrate that actual face-name matches do occur. The face-name match implies that people "live up to their given name" in their physical identity. The possibility that our name can influence our look, even to a small extent, is intriguing, suggesting the important role of social structuring in general and naming in particular in the complex interaction between the self and society. We are subject to social structuring from the minute we are born, not only by our gender, ethnicity, and socioeconomic status, but also by the simple choice that others make in giving us our name.

References

Alvergne, A., Faurie, C., & Raymond, M. (2007). Differential facial resemblance of young children to their parents: Who do children look like more? *Evolution and Human Behavior, 28,* 135–144. http://dx.doi.org/10.1016/j.evolhumbehav.2006.08.008

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

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6,* 269–278. http://dx.doi.org/10.1037/1528-3542.6.2.269

Berry, D. S. (1991). Accuracy in social perception: Contributions of facial and vocal information. *Journal of Personality and Social Psychology, 61,* 298–307. http://dx.doi.org/10.1037/0022-3514.61.2.298

Berry, D. S., & Brownlow, S. (1989). Were the physiognomists right? Personality correlates of facial babyishness. *Personality and Social Psychology Bulletin, 15,* 266–279. http://dx.doi.org/10.1177/0146167289152013

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review, 94,* 991–1013. http://dx.doi.org/10.1257/0002828042002561

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327. http://dx.doi.org/10.1111/j.2044-8295.1986.tb02199.x

Bruning, J. L., Polinko, N. K., & Buckingham, J. T. (1998). Connotative meaning of names and nicknames: Changes over twenty-seven years. *Psychological Reports, 83,* 427–430. http://dx.doi.org/10.2466/pr0.1998.83.2.427

Chen, H., Gallagher, A., & Girod, B. (2013). What's in a name? First names as facial attributes. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33,* 541–560. http://dx.doi.org/10.1006/jesp.1997.1329

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

Cohen, G., & Burke, D. M. (1993). Memory for proper names: A review. *Memory, 1,* 249–263. http://dx.doi.org/10.1080/09658219308258237

Crocker, J., Major, B., & Steele, C. (1998). Social stigma. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 504–553). Boston, MA: McGraw-Hill.

Darley, J. M., & Fazio, R. H. (1980). Expectancy Confirmation Processes Arising in the Social Interaction Sequence. *American Psychologist, 35,* 867–881. http://dx.doi.org/10.1037/0003-066X.35.10.867

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8,* 431–439. http://dx.doi.org/10.1068/p080431

Flugel, I. (1930). On the significance of names. *British Journal of Psychology, 10,* 208–213. http://dx.doi.org/10.1111/j.2044-8341.1930.tb01017.x

Fryer, R. G., & Levitt, S. D. (2004). The causes and consequences of distinctively black names. *The Quarterly Journal of Economics, 119,* 767–805. http://dx.doi.org/10.1162/0033553041502180

Ganzach, Y., & Schul, Y. (1995). The influence of quantity of information and goal framing on decision. *Acta Psychologica, 89,* 23–36. http://dx.doi.org/10.1016/0001-6918(94)00004-Z

Gordon, I., & Tanaka, J. W. (2011). Putting a name to a face: The role of name labels in the formation of face memories. *Journal of Cognitive Neuroscience, 23,* 3280–3293. http://dx.doi.org/10.1162/jocn_a_00036

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109,* 3–25. http://dx.doi.org/10.1037/0033-295X.109.1.3

Haselhuhn, M. P., Wong, E. M., Ormiston, M. E., Inesi, M. E., & Galinsky, A. D. (2014). Negotiating face-to-face: Men's facial structure predicts negotiation performance. *The Leadership Quarterly, 25,* 835–845. http://dx.doi.org/10.1016/j.leaqua.2013.12.003

Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology, 78,* 837–852. http://dx.doi.org/10.1037/0022-3514.78.5.837

Huang, G. B., Jain, V., & Learned-Miller, E. (2007, October). Unsupervised joint alignment of complex images. In Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference (pp. 1–8). IEEE.

Insaf, S. (2002). Not the same by any other name. *Journal of the American Academy of Psychoanalysis, 30,* 463–473. http://dx.doi.org/10.1521/jaap.30.3.463.21965

INSEE. (2011). Fichier des prénoms 2011 [Database of the French Given Names, produced by the National Institute of Statistics and Economic Studies]. Electronic resource.

Jones, B. C., Hahn, A. C., Fisher, C. I., Wincenciak, J., Kandrik, M., Roberts, S. C., . . . DeBruine, L. M. (2015). Facial coloration tracks changes in women's estradiol. *Psychoneuroendocrinology, 56,* 29–34. http://dx.doi.org/10.1016/j.psyneuen.2015.02.021

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103,* 54–69. http://dx.doi.org/10.1037/a0028347

Kaminski, G., Méary, D., Mermillod, M., & Gentaz, E. (2011). Is it a he or a she? Behavioral and computational approaches to sex categorization. *Attention, Perception, & Psychophysics, 73,* 1344–1349. http://dx.doi.org/10.3758/s13414-011-0139-1

Köhler, W. (1929). *Gestalt psychology*. New York, NY: Horace Liveright.

Kramer, R. S. S., & Jones, A. L. (2015). Do people's first names match their faces? *Journal of Articles in Support of the Null Hypothesis, 12,* 1–8.

Kreiborg, S., Jensen, B. L., Møller, E., & Björk, A. (1978). Craniofacial growth in a case of congenital muscular dystrophy. *American Journal of Orthodontics, 74,* 207–215. http://dx.doi.org/10.1016/0002-9416(78)90086-6

Laham, S. M., Koval, P., & Alter, A. L. (2012). The name-pronunciation effect: Why people like Mr. Smith more than Mr. Colquhoun. *Journal of Experimental Social Psychology, 48,* 752–756. http://dx.doi.org/10.1016/j.jesp.2011.12.002

Lea, M. A., Thomas, R. D., Lamkin, N. A., & Bell, A. (2007). Who do you look like? Evidence of facial stereotypes for male names. *Psychonomic Bulletin & Review, 14,* 901–907. http://dx.doi.org/10.3758/BF03194119

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE.*

Lefevre, C. E., Etchells, P. J., Howell, E. C., Clark, A. P., & Penton-Voak, I. S. (2014). Facial width-to-height ratio predicts self-reported dominance and aggression in males and females, but a measure of masculinity does not. *Biology Letters, 10,* 20140729. http://dx.doi.org/10.1098/rsbl.2014.0729

Lieberson, S., & Bell, E. O. (1992). Children's first names: An empirical study of social taste. *American Journal of Sociology, 98,* 511–554. http://dx.doi.org/10.1086/230048

Livingston, R. W., & Pearce, N. A. (2009). The teddy-bear effect: Does having a baby face benefit black chief executive officers? *Psychological Science, 20,* 1229–1236. http://dx.doi.org/10.1111/j.1467-9280.2009.02431.x

Lõhmus, M., Sundström, L. F., & Björklund, M. (2009). Dress for success: Human facial expressions are important signals of emotions. *Annales Zoologici Fennici, 46,* 75–80. http://dx.doi.org/10.5735/086.046.0108

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous-race face effect. *Psychology, Public Policy, and Law, 7,* 98–118. http://dx.doi.org/10.1037/1076-8971.7.1.98

Maglio, S. J., Rabaglia, C. D., Feder, M. A., Krehm, M., & Trope, Y. (2014). Vowel sounds in words affect mental construal and shift preferences for targets. *Journal of Experimental Psychology: General, 143,* 1082–1096. http://dx.doi.org/10.1037/a0035543

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

McLain, D. K., Setters, D., Moulton, M. P., & Pratt, A. E. (2000). Ascription of resemblance of newborns by parents and nonrelatives. *Evolution and Human Behavior, 21,* 11–23. http://dx.doi.org/10.1016/S1090-5138(99)00029-X

Mehrabian, A. (1997). Impressions created by given names. *Names, 45,* 19–33. http://dx.doi.org/10.1179/nam.1997.45.1.19

Mehrabian, A. (2001). Characteristics attributed to individuals on the basis of their first names. *Genetic, Social, and General Psychology Monographs, 127,* 59–88.

Meissner, C. A., & Brigham, J. C. (2001). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology, 15,* 603–616. http://dx.doi.org/10.1002/acp.728

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 109,* 16474–16479. http://dx.doi.org/10.1073/pnas.1211286109

Oberndorf, C. P. (1920). Reaction to personal names. *The International Journal of Psycho-Analysis, 1,* 223–230.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 105,* 11087–11092. http://dx.doi.org/10.1073/pnas.0805664105

Pantelis, P. C., van Vugt, M. K., Sekuler, R., Wilson, H. R., & Kahana, M. J. (2008). Why are some people's names easier to learn than others? The effects of face similarity on memory for face-name associations. *Memory & Cognition, 36,* 1182–1195. http://dx.doi.org/10.3758/MC.36.6.1182

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

Peterson, M. A., & Rhodes, G. E. (2003). *Perception of faces, objects, and scenes: Analytic and holistic processes*. New York, NY: Oxford University Press.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8,* 364–382. http://dx.doi.org/10.1207/s15327957pspr0804_3

Roberts, T., & Bruce, V. (1988). Feature saliency in judging the sex and familiarity of faces. *Perception, 17,* 475–481. http://dx.doi.org/10.1068/p170475

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

Sapir, E. (1929). A study of phonetic symbolism. *Journal of Experimental Psychology, 12,* 225–239. http://dx.doi.org/10.1037/h0070931

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10,* 80–83. http://dx.doi.org/10.1111/1467-9280.00111

Shoval, G., Manor, I., Nahshoni, E., Weizman, A., & Zalsman, G. (2012). Are names of children with attention deficit hyperactivity disorder more 'hyperactive'? *Psychopathology, 45,* 215–219. http://dx.doi.org/10.1159/000330767

Sinclair, S., Hardin, C. D., & Lowery, B. S. (2006). Self-stereotyping in the context of multiple social identities. *Journal of Personality and Social Psychology, 90,* 529–542. http://dx.doi.org/10.1037/0022-3514.90.4.529

Slepian, M. L., & Galinsky, A. D. (2016). The voiced pronunciation of initial phonemes predicts the gender of names. *Journal of Personality and Social Psychology, 110,* 509–527. http://dx.doi.org/10.1037/pspa0000041

Štruc, V., Pavešic, N., & Pavešic, N. (2011). Photometric normalization techniques for illumination invariance. In Y. Zhang (Ed.), *Advances in face image analysis: Techniques and technologies* (pp. 279–300). Hershey, PA: Medical Information Science. http://dx.doi.org/10.4018/978-1-61520-991-0.ch015

Swim, J. K., & Stangor, C. (Eds.). (1998). *Prejudice: The target's perspective*. New York, NY: Academic.

Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1701–1708).

Tanaka, J. W. (2001). The entry point of face recognition: Evidence for face expertise. *Journal of Experimental Psychology: General, 130,* 534–543. http://dx.doi.org/10.1037/0096-3445.130.3.534

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 46,* 225–245. http://dx.doi.org/10.1080/14640749308401045

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. In A. Kingstone & M. B. Miller (Eds.), *The year in cognitive neuroscience 2008* (pp. 208–224). Malden, MA: Blackwell. http://dx.doi.org/10.1196/annals.1440.012

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. http://dx.doi.org/10.1146/annurev-psych-113011-143831

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences, 12,* 455–460. http://dx.doi.org/10.1016/j.tics.2008.10.001

Tskhay, K. O., & Rule, N. O. (2016). People automatically extract infants' sex from faces. *Journal of Nonverbal Behavior, 40,* 247–254.

Valentine, T., Bredart, S., Lawson, R., & Ward, G. (1991). What's in a name? Access to information from people's names. *European Journal of Cognitive Psychology, 3,* 147–176. http://dx.doi.org/10.1080/09541449108406224

van 't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition, 108,* 796–803. http://dx.doi.org/10.1016/j.cognition.2008.07.002

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference (Vol. 1, pp. I–511). IEEE.

Walker, M., & Vetter, T. (2016). Changing the personality of a face: Perceived Big Two and Big Five personality factors modeled in real photographs. *Journal of Personality and Social Psychology, 110,* 609–624. http://dx.doi.org/10.1037/pspp0000064

Waynbaum, I. (1907). *La Physionomie Humaine: Son mécanisme et son rôle social* [The human face: Its mechanism and social function]. Paris, France: Alcan.

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin, 127,* 797–826. http://dx.doi.org/10.1037/0033-2909.127.6.797

Wilde, O. (1974). The picture of Dorian Gray. Oxford, UK: Oxford University Press. (Original work published 1890)

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17,* 592–598. http://dx.doi.org/10.1111/j.1467-9280.2006.01750.x

Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica, 114,* 101–114. http://dx.doi.org/10.1016/S0001-6918(03)00052-0

Yovel, G., Halsband, K., Pelleg, M., Farkash, N., Gal, B., & Goshen-Gottstein, Y. (2012). Can massive but passive exposure to faces contribute to face recognition abilities? *Journal of Experimental Psychology: Human Perception and Performance, 38,* 285–289. http://dx.doi.org/10.1037/a0027077

Zajonc, R. B. (1985). Emotion and facial efference: A theory reclaimed. *Science, 228,* 15–21. http://dx.doi.org/10.1126/science.3883492

Zajonc, R. B., Adelmann, P. K., Murphy, S. T., & Niedenthal, P. M. (1987). Convergence in the physical appearance of spouses. *Motivation and Emotion, 11,* 335–346. http://dx.doi.org/10.1007/BF00992848

Zebrowitz, L. A. (1996). Physical appearance as a basis of stereotyping. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 79–120). New York, NY: Guilford Press.

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

Zebrowitz, L. A. (2006). Finally, faces find favor. *Social Cognition, 24,* 657–701. http://dx.doi.org/10.1521/soco.2006.24.5.657

Zebrowitz, L. A. (2011). Ecological and social approaches to face perception. In A. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford handbook of face perception* (pp. 31–50). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199559053.013.0003

Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and Social Psychology Review, 1,* 204–223. http://dx.doi.org/10.1207/s15327957pspr0103_2

Zebrowitz, L. A., Collins, M. A., & Dutta, R. (1998). The relationship between appearance and personality across the life span. *Personality and Social Psychology Bulletin, 24,* 736–749. http://dx.doi.org/10.1177/0146167298247006

Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer vision–ECCV 2014* (pp. 818–833). Geneva, Switzerland: Springer International Publishing.

Appendix

Detailed Description of the Computational Procedure

Using computational methods to make sense of real-world images has been at the forefront of research in statistical learning and computer vision for several decades. In part because of the unprecedented amount of publicly available digital images, recent years have seen a surge in the predictive performance of learning algorithms on diverse tasks such as object recognition, foreground-background segmentation, face detection, pose estimation, and others. On many of these tasks, learning algorithms now perform on par with or even better than humans.

For image-classification tasks, the current state-of-the-art is a class of algorithms called Deep Convolutional Neural Networks (see for instance Taigman et al., 2014), a modern variant of the classic Artificial Neural Network algorithm. These algorithms contain a simplified model of a set of layered, interconnected, artificial "neurons." The first layers receive the image as input (where each pixel is fed into a different neuron), and the last layer generates a match. Analysis of the intermediate layers often reveals visual patterns of varying complexity useful for classification. The learning phase is responsible for determining the strength of neuronal connections between each pair of linked neurons in all layers; in this way, the network amplifies the neurons related to important facial regions and discards redundant ones. A neural network can be configured in numerous ways; for Studies 4 and 7 we used the LeNet architecture proposed by LeCun et al. (1998). This is a basic architecture, but is robust and has been shown to perform well in various image-classification tasks. Using newer high-end Neural Networks would have probably required a sample set larger than ours by an order of magnitude.

When designing the computerized experiment, our goal was to have it resemble the human-subject studies as much as possible, while still allowing us to take advantage of the benefits of a computational approach. Given this aim, we decided in each trial to present the algorithm with a facial image and two candidate names, a true name and a filler. Using only two names per trial allowed us to test and analyze all name pairs, and to have each name play both the target and filler role in different trials. Using several names per trial (as in most of our studies) would not have allowed us to compare all possible combinations, and having all names as candidate names (known as multiclass learning) would have taken us farther from the human studies than desired.

As described in Study 4, we followed a "one vs. one" approach (a standard machine-learning approach) and trained a different neural network for each possible pair of names. Each name-pair neural network was trained using a random 80%–20% train–test split of the data. The training set was further partitioned to create validation sets, used in a fivefold cross-validation process to select the neural network's metaparameters (e.g., regularization constant). Because each task consists of choosing one of two possible names, we generated balanced sample sets by constraining the number of images for each name to be roughly the same. After training our algorithm on the training set, we evaluated its accuracy on the test set, and repeated the above process for all pairs of names. This led to a meta-algorithm composed of a trained neural network for each pair of names. When the meta-algorithm is presented with input from a trial (that is, a facial image and two candidate names), it feeds the image as input to the neural network corresponding to the two names, and outputs the network's match.

High-accuracy matches typically require an immensely large image dataset, especially when training deep neural networks. Modern Deep Convolutional Neural Networks are usually trained on millions of images. Compared with these standards, our minimum of 500 images per name and our overall number of images used for training is relatively small, especially because of the difficulty of our matching task. We believe that more images and more powerful algorithms should potentially lead to better matching results.

(*Appendix continues*)

In Study 7 we used the trained neural networks from Study 4 to generate heat maps—images that describe the areas in the targets that are important for accurate matches. In practice, a neural network can generate probabilistic matches. In our case, this means that instead of matching one of two possible names, it assigns probabilities to each. In Study 5 the algorithm simply matched the name with the higher probability. For heat maps, we use the actual matching probabilities, which give a smoother, probabilistic interpretation of the algorithm's output. It is important to note that these heat maps represent the facial components that are essential for matching for the algorithm. Different algorithms (e.g., based on facial key points and not only raw pixel values) and different heat-map schemes (e.g., alterations other than local obscuring) can potentially lead to different outcomes. It would be interesting to explore the relation between our heat maps and those that can be generated for human subjects—for instance, by using eye-tracking devices.

Received May 30, 2016
Revision received December 9, 2016
Accepted December 10, 2016 ■