



# Broadening the stimulus set: Introducing the American Multiracial Faces Database

Jacqueline M. Chen<sup>1</sup> · Jasmine B. Norman<sup>1</sup> · Yeseul Nam<sup>1</sup>

Published online: 23 July 2020

© The Psychonomic Society, Inc. 2020

## Abstract

Face-based perceptions form the basis for how people behave towards each other and, hence, are central to understanding human interaction. Studying face perception requires a large and diverse set of stimuli in order to make ecologically valid, generalizable conclusions. To date, there are no publicly available databases with a substantial number of Multiracial or racially ambiguous faces. Our systematic review of the literature on Multiracial person perception documented that published studies have relied on computer-generated faces (84% of stimuli), Black-White faces (74%), and male faces (63%). We sought to address these issues, and to broaden the diversity of available face stimuli, by creating the American Multiracial Faces Database (AMFD). The AMFD is a novel collection of 110 faces with mixed-race heritage and accompanying ratings of those faces by naive observers that are freely available to academic researchers. The faces (smiling and neutral expression poses) were rated on attractiveness, emotional expression, racial ambiguity, masculinity, racial group membership(s), gender group membership(s), warmth, competence, dominance, and trustworthiness. The large majority of the AMFD faces are racially ambiguous and can pass into at least two different racial categories. These faces will be useful to researchers seeking to study Multiracial person perception as well as those looking for racially ambiguous faces in order to study categorization processes in general. Consequently, the AMFD will be useful to a broad group of researchers who are studying face perception.

**Keywords** Face stimuli · Face perception · Multiracial · Mixed-race · Impression formation

Perceptions of the human face play a unique and essential role in social interaction. Perceivers glean a wealth of information from an individual's facial appearance, from their membership in social groups, such as race and gender, to judgments of the person's trustworthiness and warmth. Although face-based perceptions range quite a bit in their actual accuracy, these perceptions form the basis for how people behave towards each other (Jaeger et al., 2019; Gunaydin et al., 2017). In addition, the traits conveyed by a person's face can predict their outcomes in important domains such as education (Williams et al., 2019), the criminal justice system (Chen et al., 2020; Eberhardt et al., 2006; Wilson & Rule, 2015, 2016), and healthcare (Mattarozzi et al., 2017). Therefore, face-based impression formation processes are of fundamental importance to understanding human interaction.

Although face perception processes have been widely studied, recent evidence suggests that they are not as well understood as previously thought. In particular, studies that examined variability by target race and perceiver race indicate that previously validated models of face perception may have limitations in their generalizability. Specifically, Todorov and colleagues' seminal work on impression formation concluded that perceivers extract facial information on two fundamental dimensions of dominance and trustworthiness (Oosterhof & Todorov, 2008; Todorov et al., 2008). Yet these dimensions received inconsistent support in a global investigation conducted by Jones et al. (2019), who replicated the previous findings when using the original statistical approach but documented a novel third dimension when using a different statistical approach. Furthermore, Xie et al. (2019) showed that the structure of face-based impressions differs by targets' race and gender. These recent findings raise questions about the generalizability of previous face perception findings to more diverse samples of targets and perceivers.

Another example of how diversity-related issues have accelerated growth in face perception research is the literature on racial categorization. Classic impression formation theories

---

✉ Jacqueline M. Chen  
Jacqueline.chen@psych.utah.edu

<sup>1</sup> Department of Psychology, University of Utah, Salt Lake City, UT 84112, USA

assumed that targets were monoracial (Brewer, 1998; Fiske & Neuberg, 1990), and theorized that racial categorizations are made quickly and efficiently, and consequently activate a host of race-related attitudes and stereotypes. By virtue of their potential membership in multiple existing racial groups (e.g., Black, White) and that these membership(s) may not be readily discerned from their appearance, Multiracial individuals challenge the assumptions of classic impression formation models (Chen & Hamilton, 2012; Norman & Chen, 2019; Parker et al., 2015) and models of intergroup relations that focus on categorization as an antecedent process to stereotyping, prejudice, and discrimination (e.g., Gaertner et al., 2000; Tajfel & Turner, 1979). Consistent with this, a growing body of work demonstrates that the perception of race, and the process of racial categorization, is a complex process driven by the integration of bottom-up cues and top-down cognitive influences (Freeman & Ambady, 2011). Thus, considering the case of Multiracial individuals is important both for the enrichment of our theories and for the generalizability of our literature to broader populations.

Until recently, Multiracial individuals have been historically underrepresented in face perception research. Over the past decade or so, researchers have begun to investigate the perception of Multiracial faces, covering a range of topics such as racial categorization, evaluation, and person memory. These studies have produced novel insights not only about how Multiracial individuals may be perceived and treated, but also about lay theories and assumptions surrounding race (see Chen, 2019; Ho et al., 2020; Pauker et al., 2018a, b). Nevertheless, the generalizability and growth of this research area has been hampered by reliance on a small and narrow set of stimuli. Published studies' reliance on a small set of stimuli amplifies the limitations of existing face stimuli, which we comprehensively review in this article.

To our knowledge, there is no publicly available database with a substantial number of faces belonging to real people who are Multiracial. Prior to making this conclusion, we conducted a thorough internet search for face databases and identified three websites with comprehensive lists of face databases (Evolved Person Perception and Cognition Lab, n.d.; Grgic & Delac, 2019; Stolier, n.d.). We searched all of the listed databases on the three websites for mention of Multiracial, and we conducted cold internet searches for any additional sets of faces belonging to Multiracial individuals. Based on this review, there was only one published database with more than one Multiracial stimulus face, the Disguised Face Database (2017), of which 6% (approx. 19) of their 325 faces were designated as identifying as Other *or* Multiracial. In other words, researchers who are interested in obtaining real Multiracial faces for research use would currently be able to find a maximum of 19 faces (probably fewer) over the internet.

We sought to address the field's need for Multiracial face stimuli by creating the American Multiracial Faces Database (AMFD), a novel collection of 110 unique faces of self-reported Multiracial individuals and accompanying ratings of those faces by naive observers. It is important to note that our aim was to provide faces of actual people with a mixed-race background, regardless of whether they are *perceived* as Multiracial. Past research indicates that there are both perceiver- and target-driven factors that inhibit the extent to which others perceive faces, even highly ambiguous ones, to be Multiracial (see Chen & Hamilton, 2012; Chen et al., 2018; Ho et al., 2020). Therefore, we sought to increase the number of real Multiracial faces without regard to how they are perceived by others. As discussed earlier, we believe that the AMFD will be useful to a broad group of researchers who are interested in studying face perception processes, and in particular those who are interested in understanding how Multiracial people are perceived by others, the latter of which we expand on in the next section.

## Stimulus-driven limitations of the published Multiracial face perception literature

One important contribution of the AMFD is that it begins to address the existing limitations of the literature on Multiracial face perception. To systematically assess these limitations, we conducted a review of the stimuli used in the published literature on Multiracial person perception.

Our review of the literature involved several steps. To be included, articles needed to claim to generate knowledge about the perception of Multiracial individuals and use face stimuli. First, we reviewed the references from two recent review articles on this literature (Chen, 2019; Pauker et al., 2018b). Next, we searched for relevant articles using Google Scholar and APA PsycArticles with all possible combinations of the terms “multiracial,” “biracial,” or “mixed,” and “face perception,” “person perception,” and “racial categorization.” Then, for the identified articles, we searched the list of articles *citing* the identified article for additional papers meeting the inclusion criteria. Finally, any author who occurred on the list of articles more than once became an internet search term to examine their publications for additional relevant articles (especially those recently published or in press). This literature review resulted in 48 articles (see Table 1).<sup>1</sup>

Next, we coded the articles for methodological details. We counted the number of stimuli used across all studies within

<sup>1</sup> Note that excluded from our review were articles on Multiracial person perception that did not use face stimuli. We also excluded research articles that used racially ambiguous face stimuli to examine social perception processes without any explicit goal to determine perceptions of Multiracial individuals. These studies were not included in our review, though research of this nature may also find our stimuli to be useful.

**Table 1.** Summary of published studies using face stimuli to study Multiracial person perception

No.	Author	Year	Study no.	Source	Real?	Type of face
1	MacLin & Malpass	2001	1	Original	No	LB
2	MacLin & Malpass	2003	2	Study 1	No	LB
3	Willadsen-Jensen & Ito	2006	Single	MacLin & Malpass (2001)	No	LB
4	Shutts & Kinzler	2007	1	Original	No	AW
5	Hutchings & Haddock	2008	2	Original	No	BW
6	Peery & Bodenhausen	2008	1	Original	No	BW
7	Willadsen-Jensen & Ito	2008	2	Study 1 + Original	No	BW
8	Pauker & Ambady	2009	Single	Original	No	BW
9	Pauker et al.	2009	1	Minear & Park (2004)	No	BW
10	Herman	2010	2	Study 1	No	BW
11	Miller, Maner, & Becker	2010	3	Study 1 + Original	No	BW
12	Halberstadt, Sherman, & Sherman	2011	4	Study 3	No	BW
13	Ho, Sidanius, Levin, & Banaji	2011	Single	Original	Yes	Not specified
14	Ito, Willadsen-Jensen, Kaye, & Park	2011	5	Original	No	BW
15	Newton, Dickter, & Gyurovski	2011	1	Original	No	AW
16	Chen & Hamilton	2012	3A	Original	No	AW, BW
17	Dickter & Kittel	2012	3B	Study 3A	No	AW, BW
18	Plaks, Malahy, Sedlins, & Shoda	2012	1	Willadsen-Jensen & Ito (2006)	No	BW
19	Rodeheffer, Hill, & Lord	2012	2	Study 1	No	BW
20	Krosch, Berntsen, Amodio, Jost, & van Bavel	2013	3	Study 1	No	BW
21	Pauker, Ambady, & Freeman	2013a	Single	Original	No	BW
22	Young, Sanchez, & Wilton	2013	1	Original	No	BW
23	Chen, Moons, Gauthier, Hamilton, & Sherman	2014	2	Study 1	No	BW
24	Gauthier et al.	2014	3	Study 1	No	BW
25	Halberstadt & Winkielman	2014	Single	Original	No	AW
			1	Peery & Bodenhausen (2008) + Original	Morphed and real	BW
			2	Peery & Bodenhausen (2008) + Original	Morphed and real	BW
			3	Original	No	BW
			4	Study 1	No	BW
			5	Study 2	No	BW
			6	Study 2	No	BW
			1	Original	No	BW
			2	Study 1	No	BW
			2	Original	No	BW
			2	Study 1	No	BW
			3	Study 1	No	BW
			Single	Original	No	BW
			Single	Original	Yes	BW
			1	Minear & Park (2004)	Morphed and real	BW
			2	Pauker, Johnson & Ambady (2013b) + Peery & Bodenhausen (2008)	Morphed and real	BW
			2	Study 1	Morphed and real	BW
			3	Study 1	Morphed and real	BW
			3	Pauker et al. (2009)	No	BW
			3	Original	No	AW

Table 1. (continued)

No.	Author	Year	Study no.	Source	Real?	Type of face
26	Wilton, Sanchez, & Giamo	2014	4	Original	No	AW
27	Chen & Ratliff	2015	1	Original + Minear & Park (2004) Pauker, Ambady, & Freeman (2013a), Chen et al. (2014)	No	AW
			2	Study 1	Yes	BW
			3	Study 1	Yes	BW
28	Ho, Roberts, & Gelman	2015	2	Original	No	BW
29	Kang, Plaks, & Remedios	2015	1	Original	No	BW
			2	Halberstadt & Winkielman (2014; Study 3)	No	BW
			3	Original	No	BW
30	Roberts & Gelman	2015	1	Lobue & Thrasher (2015)	Yes	BW
			2	Study 1	Yes	BW
31	Skinner & Nicolas	2015	1	Original	No	BW
			2	Study 1	No	BW
32	Willadsen-Jensen & Ito	2015	Single	Willadsen-Jensen & Ito (2006)	No	BW
33	Freeman, Pauker, & Sanchez	2016	1	Original	No	BW
			2	Study 1 + Chen et al. (2018)	No	BW
34	Gaither, Pauker, Slepian, & Sommer	2016	1	Pauker et al. (2009)	No	BW
			2A	Study 1	No	BW
			2B	Study 1	No	BW
			3	Study 1	No	BW
35	Roberts & Gelman	2017	Single	Langner et al. (2010), Lobue & Thrasher (2015), Kinzler & Dautel (2012), online sources	Yes	BW
36	Roberts, Leonard, Ho, & Gelman	2017a	Single	Lobue & Thrasher (2015)	Yes	BW
37	Roberts, Williams, & Gelman	2017b	1	Lobue & Thrasher (2015) + online sources	Yes	BW
			2	Study 1	Yes	BW
			3	Study 1	Yes	BW
38	Young, Sanchez, & Wilton	2017	Single	Minear & Park (2004) + Pauker et al. (2009)	Yes	BW
39	Chen, Pauker, Gaither, Hamilton, & Sherman	2018	1	Pauker, Ambady, & Freeman (2013a), Chen et al. (2014)	Yes	BW
			2	Chen et al. (2014)	Yes	BW
			3	Original	No	BW
40	Cooley, Brown-Iannuzzi, Brown, & Polikoff	2018	1	Chicago Face Database + Strohminger et al. (2016) + Minear & Park (2004) + Online	Yes	BW
			2	Study 1	Yes	BW
			3A	Study 1	Yes	BW
			3B	Study 1	Yes	BW
			4A	Study 1	Yes	BW
			4B	Study 1	Yes	BW
41	Gaither, Babbitt, & Sommers	2018a	1	Original	Yes (real person)	BW
			2	Photo of Study 1	Yes	BW
42	Gaither, Chen, Pauker, & Sommers	2018b	1	Pauker et al. (2009) + Gaither et al. (2014) + Himzman & Kelly (2013)	Morphed and real	BW
			2	Study 1	Morphed and real	BW
			3	Study 1	Morphed and real	BW
			4A	Study 1 + Original	Morphed and real	BW
			4B	Study 4A	Morphed and real	BW
43	Pauker, Carpinella, Lick, Sanchez, & Johnson	2018a	1	Original	No	AW, AB, BW

Table 1. (continued)

No.	Author	Year	Study no.	Source	Real?	Type of face
44	Wilton, Rattan, & Sanchez	2018	2	Study 1	No	AW, AB, BW
45	Gaither, Toosi, Babbitt, & Sommers	2019	2	Freeman, Pauker, & Sanchez (2016) Original	No	BW
			3	Study 2	No	BW, AW, AB
			4	Study 2	No	BW, AW, AB
			5	Study 2	No	BW, AW, AB
46	Nicolas, Skinner, & Dickter	2019	1	Original	No	BW, AW, BA
47	Small & Major	2019	2	Study 1 + Original	Morphed+	BW
			1	Minear & Park (2004)	No	BW
			2	Study 1	No	LW
48	Thorntenson, Pazda, Young, & Slepian	2019	1A	Pauker et al. (2009)	No	BW
			1B	Study 1A	No	BW
			1C	Study 1A	No	BW
			2	Study 1A	No	BW
			3	Study 1A	No	BW
			4	Study 1A	No	BW

Note. BW = Black-White, AW = Asian-White, LW = Latinx-White, AB = Asian-Black, LB = Latinx-Black, AL = Asian-Latinx

each article, which totaled 1295 unique Multiracial faces. However, this number is inflated by Herman (2010), which was an outlier among the articles because it alone used 769 faces. Because Herman's stimuli are not available for viewing or use by other researchers, it is therefore more accurate to say that the Multiracial literature is based on 526 unique stimuli. From our analysis of the stimuli (summarized in Tables 1 and 2), we gleaned three major limitations of existing research.

### Overreliance on computer-generated faces

We coded whether the stimuli used were faces of real Multiracial individuals or computer-generated faces. *Computer-generated faces* are defined as those created by morphing at least two monoracial faces together into a composite mixed-race face or by generating artificial faces with mixed-race features using software algorithms or by importing a real face to serve as the basis for the artificial one (e.g., FaceGen© faces; see Balas et al., 2018; Gaither, Chen, et al., 2018b). We found that the majority of studies examining Multiracial face perception operationalize Multiracial faces using computer-generated faces. Only 16% of the face stimuli used were real Multiracial faces ( $n = 86$  faces).

On the one hand, it is understandable that researchers would use computer-generated face stimuli. Relative to collecting real Multiracial faces, using computer-generated faces offers increased experimental control, standardization, and ease of obtaining novel stimuli. On the other hand, recent research suggests that using computer-generated faces may severely limit the ecological validity of one's findings. Most importantly, the same face viewed as a real face image or an artificial agent can elicit divergent racial categorizations (Gaither, Chen, et al., 2018b), evaluation along important dimensions (competence, aggression, health; Balas & Thrash, 2019), and even rates of recall (Balas & Pacella, 2015). Together these findings indicate that the artificiality of computer-generated faces can alter social perceptions of faces (see also Balas et al., 2018), making it problematic for researchers to rely on computer-generated faces to study how people perceive Multiracial individuals in the real world.

In sum, the Multiracial person perception literature is heavily reliant on computer-generated facial stimuli. Yet, because computer-generated stimuli may elicit different processes than real face images, it is essential that researchers gain access to a greater number of real faces of Multiracial individuals.

### Dominance of Black-White Multiracial faces

Another limitation of existing research is that the majority of published studies focus on the perception of Black-White biracial people. Of the studies we reviewed, 74% of all face stimuli used were Black-White Multiracial. Yet, according

**Table 2.** Summary of Multiracial person perception studies using original stimuli

No.	Author	Year	Study no.	No. of unique stimuli	Real	Race type	Target gender
1	MacLin & Malpass	2001	1	20	No	HB	No information
2	Willadsen-Jensen & Ito	2006	1	12	No	AW	Men only
			2	18	No	BW	Men only
3	Shutts & Kinzler	2007	1	8	No	BW	50/50
			2	8	No	BW	50/50
4	Hutchings & Haddock	2008	Single	1	No	BW	No information
5	Pauker & Ambady	2009	Single	40	Yes	AW	Varied - breakdown not specified
6	Pauker et al.	2009	1	20	No	BW	50/50
			3	20	No	BW	50/50
7	Herman	2010	Single	769	Yes	Not specified	Varied - breakdown not specified
8	Miller, Maner, & Becker	2010	5	42	No	BW	Men only
9	Halberstadt, Sherman, & Sherman	2011	1	4	No	AW	No information
10	Ho, Sidanius, Levin, & Banaji	2011	3A	2	No	AW, BW	No information
11	Newton, Dickter, & Gyurovski	2011	Single	1	No	BW	Man
12	Chen & Hamilton	2012	2	8	Morphed and real	BW	No information
			3	8	Morphed and real	AW	Women only
13	Dickter & Kittel	2012	Single	20	No	BW	Men only
14	Plaks, Malahy, Sedlins, & Shoda	2012	1	9	No	BW	No information
15	Rodeheffer, Hill, & Lord	2012	1	20	No	BW	50/50
16	Krosch, Berntsen, Amodio, Jost, & van Bavel	2013	1	10	No	BW	Men only
17	Pauker, Ambady, & Freeman	2013a	Single	20	No	BW	50/50
18	Chen, Moons, Gaither, Hamilton, & Sherman	2014	4	4	No	BW	Men only
19	Halberstadt & Winkielman	2014	3	24	No	AW	Women only
			4	5	No	AW	Women only
20	Wilton, Sanchez, & Giamo	2014	Single	4	No	AW	Men only
21	Ho, Roberts, & Gelman	2015	2	20	No	BW	Men only
22	Kang, Plaks, & Remedios	2015	1	15	No	BW	Men only
			3	3	No	BW	No information
23	Skinner & Nicolas	2015	1	3	No	BW	Men only
24	Freeman, Pauker, & Sanchez	2016	1	8	No	BW	Men only
25	Chen, Pauker, Gaither, Hamilton, & Sherman	2018	3	40	No	BW	50/50
26	Gaither, Babbit, Sommers	2018a	1	4	Yes (real person)	BW	Varied - breakdown not specified
			2	4	Yes (picture of Study 1)	BW	Varied - breakdown not specified
27	Gaither, Chen, Pauker, & Sommers	2018b	4A	40	Morphed and real	BW	13 women
28	Pauker, Carpinella, Lick, Sanchez, & Johnson	2018a	1	24	No	BW, AW, AB	50/50
29	Gaither, Toosi, Babbit, & Sommers	2019	2	3	No	BW, AW, AB	No information
30	Nicolas, Skinner, & Dickter	2019	1	26	Morphed and real	BW	50/50
			2	20	Morphed and real	BW	50/50

Note. BW = Black-White, AW = Asian-White, HW = Hispanic-White, AB =Asian-Black, HB = Hispanic-Black, AH = Asian-Hispanic

to Parker et al. (2015), the largest Multiracial subgroups are (in descending order): White-American Indian, Black-American Indian, Black-White, and Multiracial Hispanic. Among Multiracial babies (children under 1 year old), the largest Multiracial subgroups are (in descending order): White-Hispanic (42%), at least one Multiracial parent (22%), and Asian-White (14%; Livingston, 2017). Therefore, there are quite a few prominent Multiracial groups who are understudied, and researchers need a broader set of stimuli to be able to investigate questions that pertain to prominent Multiracial groups in society.

Whereas Asian-White biracial people have been investigated in several studies to date (Chen et al., 2019; Garay et al., 2019; Ho et al., 2011) and are the second most represented target group in the literature (19% of stimuli), there is scant research investigating the perception of Multiracial Hispanic faces. It should be noted that perceptions of the Hispanic/Latinx category differ from its official designation as an ethnicity (as opposed to being designated as a race). Specifically, although the U.S. Census considers Hispanic to be an “ethnicity” rather than a “race,” the majority of Hispanic individuals consider it to be both their race and ethnicity (Parker et al., 2015), and non-Hispanic perceivers utilize the Latinx category on par with the officially designated “race” categories of White, Asian, and Black (Chen et al., 2018; Nicolas et al., 2019). Moreover, over 16% of Hispanic/Latinx people identify as Multiracial (Parker et al., 2015). Thus, it is clear that both perceivers and targets consider Latinx people to be relevant to Multiracial issues. While there are some studies examining perceptions of Latinx individuals’ race (Ma et al., 2018; Sanchez & Chavez, 2010; Sanchez et al., 2012; Wilton et al., 2013; Young et al., 2016), we did not find any studies that examine Multiracial person perception with Latinx face stimuli. The one exception is a study that demonstrated that the perception of racially ambiguous faces was Latinx if their hairstyle was stereotypical of the category (MacLin & Malpass, 2001), but this work focused on hairstyle as a cue (showing that the same faces could be reliably categorized as Black when they had stereotypically Black hair). In addition, MacLin and Malpass’ study appeared to exclusively use male, computer-generated faces.

### Overrepresentation of male faces

Another issue that emerged from our literature review was the overrepresentation of male faces. Approximately 39% of studies purporting to examine perceptions of Multiracial faces relied *exclusively* on male faces, with several studies failing to explicitly describe the gender of their targets. Of the studies that did describe the gender breakdown of their stimuli, only 37% of the face stimuli used were female. While not unique to this research area, the androcentric bias in psychological research is problematic for many reasons (see Bailey et al.,

2019, for a deeper discussion of these issues). One important concern is that different perceptions of targets’ race and gender can have an interactive effect on social perception (Carpinella et al., 2015; Cole, 2009; Ghavami & Peplau, 2013; Goff et al., 2008; Johnson et al., 2012), and therefore research findings that rely on male targets may not generalize to female targets of the same race. Furthermore, androcentric biases frequently render women of color invisible (Purdie-Vaughns & Eibach, 2008; Sesko & Biernat, 2010; Stroessner, 1996), making them an understudied population across several strands of literature in psychology. Therefore, it is imperative that researchers have stimulus sets that afford them the ability to test the generalizability of their findings across target gender.

In sum, we have documented several issues that negatively impact the generalizability and validity of current findings on Multiracial face perception. These issues will only persist without the availability of new stimulus materials.

## Overview of current research

Our goal was to provide a database of face stimuli that would facilitate researchers’ ability to conduct face perception studies that can enable novel research questions, test the generalizability of published results from the impression formation and Multiracial perception literatures, and combat existing biases in the published literature. To this end, we collected photographs of the faces of real Multiracial people in a laboratory with a standardized background and lighting, and we pretested the photographs on several fundamental dimensions of impression formation (trustworthiness, dominance, warmth, competence) as well as on attributes of regular interest to researchers in face perception (e.g., attractiveness). We present the American Multiracial Face Database, a collection of 110 faces of real Multiracial people. The AMFD consists of predominantly Asian-White and Latinx-White Multiracial faces, the majority of which are female.

## Method

### Database development

#### Recruitment of Multiracial volunteers

Individuals were recruited from a large public university in the United States. They participated in the study in exchange for partial course credit. The study ad was posted online and distributed via email to several large lecture courses in the Psychology Department. The ads specified that individuals were eligible to participate if they identified as having a mixed-race background. One potential concern is that the

study ad disproportionately attracted strongly identified Multiracial individuals to the study. However, this is very unlikely because approximately 40% of all participants who signed up for the study reported that they did not have mixed race backgrounds in the study questionnaire. Therefore, it seems more likely that the ads did not effectively emphasize the participation requirements and cast a broader net than was intended. For this database, we only included participants who reported multiple racial ancestries<sup>2</sup> and who consented to having their photograph used for research purposes, resulting in a total of 110 volunteers.

Please note that we informed volunteers that their photographs would be shared with others for research use without any accompanying personal information. Therefore, the database does not report volunteers' self-reported racial background and identities, and we have collapsed across backgrounds that were not shared by at least four people in the reporting of this research. Of our 110 volunteers ( $M_{age} = 20.73$ ,  $SD = 2.42$ ; 90 women, 20 men), 81 self-reported a racial background consisting of two racial groups and 29 self-reported a racial background consisting of three or more racial groups. The sample of volunteers included approximately 33% Asian/White, 22% Latinx/White, 11% Asian/Latinx, 6% White/Middle Eastern, 5% Black/White, and 5% Asian/Middle Eastern volunteers. Approximately 18% of volunteers selected other racial backgrounds.<sup>3</sup>

## Procedure and measures

Photographs were taken with a Canon PowerShot capable of producing 16.0-megapixel images. Volunteers were positioned standing in front of a white background and were photographed by the experimenter twice, once with a neutral expression and once smiling. Volunteers viewed the photographs on the digital camera screen for approval or an additional opportunity to retake them. They then read and signed a photo release form permitting the use of the photographs for academic research and educational purposes. Volunteers could also elect not to sign the photo release form, and 21 volunteers (16% of the initial sample) did not sign it and are not included in the final sample ( $n = 110$ ). Finally, volunteers

<sup>2</sup> Participants were considered to have mixed-race heritage if their self-description of heritage included races or ethnicities that fell under two or more of the following categories: Asian, Black, Latinx, Middle Eastern, Native American, and White. This operationalization of mixed-race heritage is consistent with how the U.S. government defines people as "multiracial" based on their Census responses, with two exceptions: 1) we consider Latinx to be a racial group rather than an ethnicity (for reasons specified in the Introduction), and 2) we consider Middle Eastern to be a distinct race separate from White (consistent with disparate outcomes and treatment of Middle Eastern people vs. White people in the West; e.g., Kteily et al., 2014).

<sup>3</sup> Although the heritage information is not included in the publicly posted data set, this information is available. Researchers who are interested in conducting analyses using the volunteers' self-reported racial background and/or identities may contact the corresponding author about their inquiry.

completed survey measures (see Norman & Chen, 2019) and demographic information. Racial background was assessed by asking volunteers "What race(s) are part of your heritage?". Volunteers could select all that apply from the following response options: Asian, Black, Latino, Middle Eastern, White, Other.

## Standardization of stimuli

Using Adobe Photoshop (v2017.0.1), all photographs were resized to  $2444 \times 1718$  pixels (400 ppi), such that the target face and core facial features were approximately centered in the screen. When possible, head tilt was reduced by rotating photographs prior to cropping. Photographs were then cropped in a headshot format using the rectangular cropping tool.<sup>4</sup> Adobe Photoshop's layer tools were used to place the face on a standardized white background. Example stimuli are displayed in Fig. 1.

## Obtaining stimulus ratings

### Raters

We recruited participants to make ratings of the AMFD faces in exchange for course credit or money. Student participants were recruited from a different large public university from the one at which Multiracial volunteers were recruited. Additional participants were recruited from Amazon's Mechanical Turk (MTurk). Our sample goal was set to obtain approximately 50 raters per face for each attribute. We determined this goal based on previous research indicating that face-based ratings become stable after 45 independent observations, and much earlier depending on the particular attribute (Hehman et al., 2018). We also sought to obtain a sample that was racially diverse to maximize the representativeness of the ratings for each face.

Four hundred ninety-eight students (23.5% of the sample) and 1625 MTurk workers (76.5% of the sample) rated the faces ( $M_{age} = 31.68$ ,  $SD = 11.29$ ). The sample included 949 men (44.7%), 1133 women (53.4%), 12 other-identified genders (.6%), and 29 declined to answer (1.4%). According to their self-reported race, there were 875 White Americans, 375 Black Americans, 355 Latino/Latina Americans, and 304 Asian Americans. The remainder of the sample was 143 Multiracial, 46 Other-identified, 9 Middle-Eastern, and 32 people who did not specify their race. Each face was rated by a random selection of participants, predominantly by

<sup>4</sup> Due to variation in how participants held the white board with their ID number on it when the photograph was taken, it was occasionally necessary to crop at the base of the neck or bottom of the face rather than the shoulders. When possible, we standardized the facial position such that the tip of the nose was centered, and the distance from the top of the image to the pupils was approximately 1.5–2 inches.



Example of faces from American Multiracial Face Database with most common racial and gender categorizations.



1072 - Neutral

Racial categorizations: Latina (58% of raters);  
Multiracial (21% of raters); Black (11% of raters)

Passes as: Latina, Multiracial

98% of raters indicated female gender categorization.

Total # of raters: 62



1108 - Smiling

Racial categorization: Asian (67%), Multiracial (12%)

Passes as: Asian, Latina, Multiracial

100% of raters indicated female gender categorization.

Total # of raters: 49



1102 - Neutral

Racial categorization: Asian (37% of raters); Multiracial  
(33% of raters); Latino (26% of raters)

Passes as: Asian, Latino, Multiracial, White

100% of raters indicated male gender categorization.

Total # of raters: 43



1035 - Smiling

Racial categorization: Asian (35%), Latino (33%),  
Multiracial (18%), Black (9%)

Passes as: Asian, Latino, Multiracial

100% of raters indicated male gender categorization.

Total # of raters: 55

**Fig. 1.** Example of faces from the American Multiracial Faces Database with the most common racial and gender categorizations

MTurk workers (50–85% of raters were MTurkers for any given face).

### Procedure

Student participants completed the survey in a laboratory, and MTurk participants completed it online. The survey was hosted on Qualtrics ([www.Qualtrics.com](http://www.Qualtrics.com)). After providing informed consent, participants rated a random subset of 5–8 faces from the AMFD (the number of faces depended on the advertised length of the survey opportunity and compensation rate). Participants were randomly assigned to rate only neutral photographs or only smiling photographs.

For participants recruited through MTurk, we included an attention check embedded in the demographic items (“We would like to ensure participants read all questions. Thank you in advance for doing so, and please choose slightly agree

to demonstrate this.”). Of our original 1791 MTurk raters, 9.3% ( $n = 166$ ) answered the attention check incorrectly and were excluded from the norming data (final  $N = 1625$ ).

Each face was rated by approximately 50 raters on average (range: 36 to 71 raters depending on the particular face and attribute being rated). The modal number of raters for each face was 49–51 depending on the attribute.

### Measures

For each face, participants rated dominance, trustworthiness, competence, warmth, attractiveness, affective valence of expression, how genuine the person’s smile appeared (if smiling), racial ambiguity, and masculinity/femininity. Participants were additionally asked to provide a racial categorization, gender categorization, and ratings of how prototypical each face was of Asian, Black, Latinx, Middle Eastern, Multiracial, and White racial categories.

The racial categorization question used a categorical response scale (e.g., “What race is this person?”, with response options including Asian, Black, Multiracial, etc.). The gender categorization question also used a categorical response scale (options were Male, Female, and Other). Reliability for these two items was assessed in terms of percentage agreement across participants. See Fig. 1 for examples.

The remainder of the ratings (e.g., attractiveness, affective expression) used a Likert-response scale from 1 to 7. To assess prototypicality, raters were given the following instructions adapted from Ma et al. (2015): “People often differ in how much their physical features resemble features of various racial groups. Please rate how typical this person is in regards to the following groups,” and were asked to rate each category on a scale from 1 (*very atypical*) to 7 (*very typical*). Perceived traits were assessed by asking participants to indicate how “dominant,” “trustworthy,” “smart” (proxy for competence), and “warm” the person in the photograph appeared on a scale from 1 (*not at all*) to 7 (*extremely*). Racial ambiguity was assessed on the same rating scale (i.e., 1: *not at all ambiguous* to 7: *extremely ambiguous*). Smile ratings were assessed from 1 (*not at all genuine*) to 7 (*extremely genuine*) and included a response option labeled “This person is not smiling.” The following response scales were used for other ratings: attractiveness (1: *very unattractive* to 7: *very attractive*), affective valence of expression (1: *very negative* to 7: *very positive*), masculinity/femininity (1: *extremely masculine* to 7: *extremely feminine*).

To assess reliability across participants on these ratings, we used cross-classified multilevel models to compute intraclass correlation coefficients (ICCs). Because ratings are clustered by rater and target face, we can estimate the percentage of variance in the ratings attributable to rater characteristics, face characteristics, and their interaction (see Hehman et al., 2017; Xie et al., 2019). Because we are primarily focused on reporting the reliability indices that inform researchers about the extent to which ratings are driven by characteristics of the target faces, we report the ICC that represents the percentage of variance in ratings that comes from between-target characteristics (i.e., the measure of “consensus” across perceivers; Kenny & West, 2010). ICCs are contained in Table 3.

The data and codebook files are posted online (<https://osf.io/qsdrp/>). The codebook describes the scale and item wording for each rating and includes the number of raters that contributed to the average rating for each attribute.

## Results

The database included 110 unique individuals who, with one exception, each posed with a neutral expression and a smiling expression. The total number of photographs is 219 (109 neutral, 110 smiling). We confirmed that the smiling faces ( $M =$

5.45,  $SD = .61$ ) were rated as having more positive affective expressions compared to the neutral faces ( $M = 3.96$ ,  $SD = .67$ ),  $t(217) = -17.21$ ,  $p < .001$ ,  $d = -2.34$ .

With respect to gender, the faces were unambiguous; the highest proportion of raters classifying any face as “other” than a man or a woman was .04. Ninety faces were reliably categorized as women, and 20 were reliably categorized as men. Ratings of femininity were strongly predictive of gender categorizations (with proportion of female categorizations,  $r(217) = .94$ ,  $p < .001$ ). Table 3 reports the descriptive statistics for each rating dimension collapsed across all stimuli in the AMFD.

### Racial appearance of AMFD faces

We used the neutral expression faces to conduct the following analyses, in light of the fact that facial affect can influence perceptions of race (e.g., Hugenberg & Bodenhausen, 2004).

### Level of racial ambiguity of faces

We first examined the level of consensus in categorizations of the faces. For each face, we determined a) the most common racial categorization and b) the proportion of participants who agreed on that categorization. The bottom quartile of AMFD faces yielded less than 46% agreement among raters. The bottom half of faces yielded less than 62% agreement. The third quartile yielded less than 77% agreement. Only the top quartile of faces yielded greater than 77% agreement among raters on what race the faces were.

We conducted a one-way ANOVA to test whether the quartiles had significantly different levels of racial ambiguity. There was a main effect of quartile,  $F(3,105) = 35.95$ ,  $p < .001$ ,  $\eta_p^2 = .51$ . Follow-up comparisons with a Bonferroni adjustment revealed that the top quartile of faces ( $M = 3.04$ ,  $SD = .51$ ) was significantly less ambiguous than the lower three quartiles,  $ps < .001$ . However, the bottom three quartiles ( $M_s = 3.97, 3.85, 3.76$ ,  $SD_s = .27, .38, .26$ , respectively) were not significantly different from one another,  $ps > .21$ . Based on this finding, we collapsed the lower three quartiles into one group (“ambiguous-looking faces”; 81 faces) and compared it with the top quartile (“prototypical-looking faces”; 28 faces). The average consensus level for racial categorizations of the ambiguous-looking Multiracial faces was .54 ( $SD = .13$ ), compared with a consensus average of .89 for the prototypical-looking Multiracial faces ( $SD = .07$ ).

Of the prototypical-looking faces ( $N = 28$ ; 20 women, 8 men), 10 were categorized as Asian, 7 as White, 6 as Black, 4 as Latinx, and 1 as Middle Eastern by at least 77% of the raters.

Of the ambiguous-looking faces ( $N = 81$ ; 69 women, 12 men), we also examined the most common categorization of each face. The most common categorizations of these faces

**Table 3.** Descriptive statistics and target ICC for each attribute rated

Rating dimension	Mean (SD)	Range of mean ratings (Min-Max)	Target ICC (Neutral, Smiling)
Attractiveness	4.51 (0.59)	2.57–5.92	.17, .17
Dominance	3.20 (0.55)	2.05–4.62	.09, .07
Expression (neg to pos)	4.71 (0.98)	2.57–6.49	.28, .28
Smile sincerity	3.79 (1.50)	1.26–6.33	.18, .23
Gender (masc to fem)	4.91 (1.17)	1.98–6.37	.54, .59
Racial ambiguity	3.67 (0.51)	1.90–4.52	.09, .09
Smart	4.67 (0.45)	3.46–5.62	.08, .08
Trustworthy	4.38 (0.55)	2.78–5.42	.08, .08
Warmth	4.30 (0.89)	2.38–5.92	.17, .15
Asian prototypicality	3.41 (1.51)	1.38–6.57	.50, .50
Black prototypicality	2.46 (1.17)	1.30–6.75	.43, .42
Latinx prototypicality	4.42 (1.09)	2.07–6.33	.32, .31
Middle Eastern prototypicality	3.14 (0.75)	1.70–5.94	.15, .15
Multiracial prototypicality	4.31 (0.62)	2.10–5.69	.11, .10
White prototypicality	3.81 (1.17)	1.63–6.85	.35, .34

were: 45 Latinx, 14 Asian, 14 White, 3 Multiracial, 1 Middle Eastern, and 1 Black. Three faces categorized as Latinx and as Multiracial as roughly equal proportions. These findings should be interpreted with caution, given in the low levels of agreement overall.

### Racial passing

Since the Multiracial faces were fairly ambiguous, analysis of their most common categorization does not reveal all the ways they can be used. To better identify how these faces can be used, we examined what races the faces could “pass” as, which we defined as when faces as at or above the midpoint on the prototypicality questions for each race. On average, the faces could pass into 2.35 ( $SD = .86$ ) different racial categories. The greatest number of faces ( $n = 42$ ) could pass into three categories, followed by two categories ( $n = 39$ ), one category ( $n = 20$ ), and four categories ( $n = 8$ ). Being able to pass as Latinx and as Multiracial were significantly likely to co-occur,  $\chi^2(1) = 39.89, p < .001$ . Passing as Latinx and as Middle Eastern were also significantly likely to co-occur,  $\chi^2(1) = 5.15, p = .03$ .

We also examined into which categories faces were able to pass with an inclusive count (i.e., one face can be in multiple of the following groups). There were 87 faces who could pass as Multiracial (75 women), 72 as Latinx (60 women), 42 as White (35 women), 31 as Asian (24 women), 14 as Middle Eastern (all women), and 10 as Black (7 women).

### Perceived race of faces by heritage

Next, we determined the link between self-reported heritage and perceived race. We conducted a series of independent-samples  $t$  tests to determine whether Multiracial people with a given heritage looked more prototypical of that race than Multiracial people without that heritage. The results, displayed in Table 4, showed clearly in all cases that people with a given heritage looked more prototypical of that race than people without that heritage.

To examine whether the AMFD faces differed in their level of ambiguity compared to monoracial faces, we tested whether the AMFD faces were less frequently categorized into their corresponding heritage groups compared to monoracial faces from the Chicago Face Database (CFD; Ma et al., 2015). Using the racial categorization data that we collected for AMFD faces and the publicly available data from the CFD, we conducted independent-samples  $t$  tests comparing the proportion of raters who categorized the faces into a given racial category. We compared AMFD faces with Asian, Black, White, and Latinx heritage and CFD Asian, Black, White, and Latinx faces, respectively. As expected, AMFD faces were less consensually categorized than CFD faces in each comparison. Specifically, AMFD part-Asian faces were categorized as Asian by 36% of raters on average ( $M = .36, SD = .33$ ), whereas CFD monoracial Asian faces were categorized as Asian by 77% of raters on average ( $M = .77; SD = .25$ ),  $t(230) = 10.56, p < .001, d = 1.39$ . AMFD part-Black faces ( $M = .42, SD = .42$ ) were categorized as Black by significantly fewer raters than CFD monoracial Black faces were ( $M = .87, SD = .23$ ),  $t(247) = 9.26, p < .001, d = 1.18$ . AMFD part-White

**Table 4.** Perceived racial prototypicality by racial heritage

	No heritage ( <i>M</i> , <i>SD</i> )	Heritage ( <i>M</i> , <i>SD</i> )	<i>p</i>	<i>d</i>
Asian prototypicality	2.38 (0.56)	4.27 (1.52)	< .001	-1.63
Black prototypicality	2.18 (0.48)	4.31 (2.01)	< .001	-1.75
Latinx prototypicality	3.94 (1.02)	4.90 (0.94)	< .001	-0.98
Middle Eastern prototypicality	3.07 (0.68)	3.74 (0.87)	< .001	-0.77
White prototypicality	3.15 (0.84)	4.09 (1.20)	< .001	-0.82

faces ( $M = .25$ ,  $SD = .30$ ) were categorized as White by significantly fewer raters than CFD monoracial White faces were ( $M = .89$ ,  $SD = .19$ ),  $t(335) = 23.87$ ,  $p < .001$ ,  $d = 2.61$ . Finally, AMFD part-Latinx faces ( $M = .43$ ,  $SD = .25$ ) were categorized as Latinx by significantly fewer raters than CFD monoracial Latinx faces were ( $M = .50$ ,  $SD = .25$ ),  $t(226) = 2.11$ ,  $p = .04$ ,  $d = 0.28$ . The difference between AMFD and CFD faces with Latinx heritage was the smallest, but this is consistent with our finding that being perceived as Latinx and as Multiracial are likely to co-occur. These comparative analyses further validate the AMFD as contributing face stimuli that are distinctly more ambiguous than those offered by an existing high-quality, monoracial database.

We then examined a few subsamples of the AMFD faces that may be of direct interest to researchers who seek to investigate perceptions of Multiracial people and racial minorities more generally.

**Asian-White faces (n = 35; 30 women)** We examined how frequently these faces were categorized as Asian versus as White. If the difference in the two rates of categorization was under 10%, we considered the faces as appearing equally Asian and White. By this metric, there were 19 Asian-looking faces, 10 ambiguous faces, and 6 White-looking faces.

It was also important to examine overall categorization rates of these faces, given that raters were presented with all possible racial categories when they were asked to classify the targets by race. We examined the most common categorization of each face for the two face groups (prototypical, ambiguous). Only 7 faces were prototypical-looking, and they were all categorized as Asian more than anything else. Among the 28 ambiguous-looking faces, the predominant categorizations were: 15 Latinx, 9 Asian, 2 Multiracial, and 1 White. One face had a tie for most frequent categorization between White and Multiracial.

Examining all of the faces using the prototypicality ratings as a measure of “passing” (as we did in the previous subsection), there were 22 faces that could pass as Asian, 11 as White, 30 as Multiracial, 21 as Latinx, 2 as Middle Eastern, and 1 as Black.

**Latinx-White faces (n = 24; 19 women)** We examined how frequently these faces were categorized as Latinx vs. White. There were 13 faces that were categorized more as Latinx, 9 categorized more as White, and 2 ambiguous faces.

Six faces were prototypical-looking (3 White-looking, 3 Latinx-looking). The 18 ambiguous faces were categorized as: Latinx (10) and White (7). One face was equally likely to be categorized as Latinx and Multiracial.

Examining all the faces for what they could “pass” as, 19 were at least somewhat prototypical of Latinx, 18 of Multiracial, 15 of White, and 2 of Middle Eastern.

**Part-Latinx faces (n = 58; 49 women)** We also analyzed all faces of volunteers who reported having Latinx heritage. Of these, 36 were categorized most frequently as Latinx, 11 as White, 5 as Asian, and 4 as Black. Two faces did not have one consensus categorization; one was categorized as Multiracial and Latinx, and the other was categorized as Black, Latinx, and Multiracial.

**Part-Middle Eastern faces (n = 24; 21 women)** We analyzed all faces of volunteers who reported having Middle Eastern heritage. Of these, 12 were categorized as Latinx, 8 as White, 2 as Middle Eastern, and 2 as Asian. Compared to the other analyses by subgroup heritage, Middle Eastern faces appeared to have the lowest correspondence from heritage to perception by others.

**Part-Black faces (n = 17; 12 women)** We analyzed all faces of volunteers who reported having Black heritage. Of these, 7 were categorized as Black, 4 as Latinx, 2 as Asian, 2 as White, and 1 as Multiracial. One face was equally likely to be categorized as Black and Latinx.

## Cues to racial categorization

We have argued that the AMFD will be useful for investigating fundamental questions about the perception of race. To examine what traits and attributes were generally associated with the perception of particular racial categories, we tested which appearance-based cues correlated with particular racial categorizations of faces (Table 5) and perceived racial prototypicality (Table 6). From these analyses, a few noteworthy findings emerged.

Importantly, perceived ambiguity strongly predicted the proportion of raters who categorized the faces as Latinx,  $r(107) = .56$ ,  $p < .001$ , and Multiracial,  $r(107) = .60$ ,  $p < .001$ . Perceived racial ambiguity also strongly predicted perceived prototypicality of Latinx,  $r(107) = .74$ ,  $p < .001$ , and Multiracial categories,  $r(107) = .79$ ,  $p < .001$ . These correlations provide strong evidence that Latinx categorizations should be interpreted as markers of high racial ambiguity.

**Table 5.** Correlations between proportion of racial categorizations and other attributes among AMFD neutral expression faces

	Asian	Black	Latinx	Middle Eastern	Multiracial	White
Dominance	-.11	.17 <sup>+</sup>	.09	-.13	-.07	-.02
Trustworthiness	.22*	.23*	-.22*	.07	-.12	-.23*
Smart	.39***	.11	-.36***	.11	-.07	-.21*
Warm	-.02	.22*	-.11	.02	-.12	-.01
Attractive	.20*	-.02	-.01	-.05	-.10	-.13
Affective expression	.07	.17 <sup>+</sup>	-.16 <sup>+</sup>	-.04	-.14	.07
Smile sincerity	-.11	.20*	-.14	-.02	-.10	.14
Ambiguity	-.17 <sup>+</sup>	-.34***	.56***	.23*	.60***	-.42***

Note. <sup>+</sup> $p < .09$ , \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ . Attributes were omitted from this table if they did not have significant or marginal correlations with any racial categorization variable

Other correlations, displayed in Tables 5 and 6, were consistent with racial stereotypes. Perceived Asian appearance (overall categorizations and prototypicality ratings) was associated with greater perceived intelligence,  $ps < .001$ . Perceived Latinx appearance was associated with lower perceived intelligence,  $ps < .01$ .

Finally, there were a few correlations that were not anticipated, but appeared to be reliable across indices of both racial categorization and prototypicality. Perceived Asian appearance was associated with greater perceived attractiveness,  $ps < .05$ . Perceived White appearance was correlated with lower perceived trustworthiness and lower perceived intelligence,  $ps < .05$ .

## Discussion

Face perception is a central process in social interaction. Researchers who are interested in studying face perception, and related impression formation processes, frequently rely on convenience samples of stimuli. While there are a number of high-quality face databases available to researchers, we found only one database with any number of Multiracial faces (approx.  $n = 19$ ) available to researchers. Having greater

numbers of stimuli, in particular those that add to the heterogeneity of existing stimulus sets, increases our field's ability to produce generalizable results and empirically address a broader set of issues. We present the American Multiracial Faces Database, a collection of 110 unique Multiracial faces with accompanying ratings by a diverse sample of participants. We provide neutral expression and smiling versions of faces in order for researchers to have additional flexibility in their research questions and methods.

Based on our systematic review of the Multiracial person literature, we contend that the Multiracial person perception research area is one that will benefit immensely from this database. In particular, the AMFD addresses the major limitations of available Multiracial faces that we identified; specifically, that previous research has relied heavily on computer-generated, Black-White, and male faces. The AMFD directly addresses these limitations by substantially increasing the number of real faces, a large number from Asian and Latinx backgrounds, and who are predominantly women. We also highlighted several subsets of the AMFD stimuli that can help to accelerate research on underrepresented groups. In particular, the AMFD may facilitate a growing number of studies on: 1) biracial people outside of the Black-White binary (e.g., Asian-White, Latinx-White), 2) perceptions of part-Black

**Table 6.** Correlations between racial prototypicality ratings and other attributes among AMFD neutral expression faces

	Asian	Black	Latinx	Middle Eastern	Multiracial	White
Dominance	-.12	.17 <sup>+</sup>	.04	-.12	.03	-.03
Trustworthiness	.20*	.17 <sup>+</sup>	-.16 <sup>+</sup>	-.08	.01	-.26**
Smart	.40***	.03	-.30**	.07	.02	-.19*
Attractive	.22*	-.08	.03	.02	.10	-.07
Smile sincerity	-.13	.10	-.11	-.09	-.16 <sup>+</sup>	.09
Ambiguity	.04	-.09	.74***	.58***	.79***	-.17 <sup>+</sup>

Note. <sup>+</sup> $p < .09$ , \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ . Attributes were omitted from this table if they did not have significant or marginal correlations with any racial categorization variable

individuals (not just Black-White biracials), 3) perceptions of Latinx-looking and part-Latinx individuals, and 4) perceptions of Middle Eastern individuals, a group that faces quite a bit of prejudice and discrimination (e.g., French et al., 2013) yet remains seriously understudied in the face perception literature.

The three-quarter majority of the AMFD faces are quite racially ambiguous. These faces will be useful to researchers who investigate the impact of racial ambiguity on person perception, as well as to those who use racially ambiguous stimuli to study the impact of top-down processes on perception (e.g., Hugenberg & Bodenhausen, 2004; Freeman & Ambady, 2011). Furthermore, this database provides faces that are able to pass into multiple racial categories, especially Multiracial ( $n = 87$ ), Latinx ( $n = 72$ ), and White ( $n = 45$ ). Relying exclusively on the prototypical-looking faces from other face databases, it would not be possible for researchers to present the same target and randomly assign participants to learn that the target is a particular race (e.g., presenting the same face with different racial labels). The AMFD faces that have several different “passing” options will enable researchers to hold a target’s appearance constant while manipulating information about their racial group membership. Therefore, the present research provides stimuli that will facilitate researchers’ ability to conduct highly controlled experiments that contribute to multiple strands of literature on social perception.

In addition to the methodological contribution of the AMFD, our analyses provided insights that raise important theoretical points. In particular, we found that one-quarter minority of AMFD faces were not perceived to be racially ambiguous. This is an important observation because, as highlighted in our review, many studies assume that Multiracial individuals are racially ambiguous and operationalize “Multiracial faces” as “racially ambiguous faces.” Yet our analysis of the perceived racial ambiguity of the AMFD faces indicates that having mixed-race heritage is not equivalent to being racially ambiguous. Therefore, if researchers continue to operationalize Multiracial as racial ambiguity, their conclusions may not apply to a substantial portion of Multiracial individuals. Of course, additional research, with an even larger set of real Multiracial stimuli, is needed to systematically investigate the relationship between mixed-race heritage and appearance. We note that it is only with more stimulus sets like the AMFD that the field will be able to obtain a representative sample of Multiracial faces from which to draw ecologically valid conclusions about how they are perceived by others.

The issue of racial heritage versus appearance highlights the broader point that researcher assumptions and biases self-perpetuate based on methodological choices that they make. Dunham and Olson (2016) advanced a similar argument: “... the reliance on discrete categories in our methods has a tendency to reify or reinforce those categories within our

theories” (p. 646). For instance, assuming that Multiracial individuals are racially ambiguous reifies existing monoracial categories by increasing the perceptual distance between them and Multiracial individuals. Establishing that Multiracial and monoracial individuals overlap in perceptual face space raises the point that race might be better thought of as a multidimensional continuum rather than a series of discrete categories.

Additional insights were gained from our analyses of the attribute ratings of AMFD faces to determine what cues were related to the perception of race. First, recent seminal work by Xie et al. (2019) found that face perceptions themselves vary in the extent to which they are determined by the targets’ appearance, the perceivers, and their interaction. In this work, an inspection of our target-ICC indices showed that the highest level of rater consensus was obtained for ratings of masculinity/femininity and, to a lesser extent, for Asian and Black prototypicality. The lowest ICC indices by far were for perceived ambiguity and Multiracial prototypicality, which is consistent with previous research arguing that the Multiracial category is not well-developed among most perceivers (and therefore does not have a clear prototype; Chen & Hamilton, 2012).

A second significant finding was that Latinx categorization and Multiracial categorization were both highly correlated with perceived racial ambiguity. That Multiracial categorizations increase with racial ambiguity has been shown previously (e.g., Chen et al., 2014). Yet this study established that perceptions of Multiracial and Latinx categories are heavily overlapping due to this shared physical cue. This result clarifies why Multiracial categorizations are fairly low in this data set as well as in previous work (Chen & Hamilton 2012; Chen et al., 2014). Specifically, our results indicate that perceivers are likely to classify an ambiguous face as Latinx, a well-established and information-rich category, rather than as Multiracial, a broad and novel category (see also Pauker et al., 2018a). In fact, our findings dovetail nicely with those of Chen et al. (2018) and Nicolas et al. (2019), who both found that the most common categorization of racially ambiguous faces was Latinx, not Multiracial, using two independent stimulus sets. This finding also helps clarify why researchers have previously struggled to identify the physical markers that underlie Latinx categorizations (Ma et al., 2018).

We also documented several correlations that replicated previous work on stereotyping. Namely, Asian-looking people were perceived as competent, and Latinx-looking people were perceived as low in competence (Fiske et al., 2002; Jimeno-Ingrum et al., 2009). Our findings also documented some negative biases against White-looking individuals, who were perceived as lower in trustworthiness and competence. This result could indicate some prejudice against White-looking individuals, or it could result from raters’ reduced self-monitoring when they evaluated White-looking faces compared to when they evaluated faces that looked non-

White, and is worthy of future investigation. In summary, the AMFD faces and ratings provides both practical and theoretical contributions to the fields of face perception, impression formation, and intergroup relations.

The AMFD does have some limitations. The quality of our images is lower than some other databases (Ma et al., 2015; Strohminger et al., 2016), which provide extremely high-definition face photographs. However, it should be noted that researchers are still using lower-quality images of older databases; for example, the Radboud Face Database (Langner et al., 2010) includes photos with resolutions between 10 and 12 megapixels (compared to 16 megapixels for the AMFD images) and was cited by 206 articles in 2019 alone. Also, given that a large proportion of research is now conducted through online surveys with image size limitations, we believe that the AMFD images will be useful to researchers across the behavioral sciences for many years to come.

Another limitation of the AMFD is that it is demographically skewed in terms of age and gender. The AMFD only includes young adults, reflecting the fact that the Multiracial population is a young population (the median age of Multiracial Americans is 19, compared to 38 for monoracial Americans; Parker et al., 2015). Nonetheless, it will be important for future research to collect faces of Multiracial individuals as they age. Finally, our database is predominantly biracial, containing the faces of people of two races, and predominantly Asian-White and Latinx-White. While the AMFD adds substantial diversity to existing face databases, it is by no means representative of the Multiracial population. In the future, we hope that researchers will curate additional databases that increase the representation of other types of Multiracial individuals, including those who have more than two racial ancestries. With respect to the gender composition of the AMFD, it reflects the fact that volunteers for psychology studies are predominantly women. Although the gender imbalance among faces is not ideal, we believe that it presents a valuable resource in particular because this imbalance is in the opposite direction of the documented bias in the Multiracial perception literature, which currently relies on exclusively or predominantly male face stimulus sets. We hope that the availability of the AMFD will help researchers address this bias in their future work, and that it may be possible for researchers to use a combination of stimuli from the AMFD and other databases with more male faces to present gender-even stimulus sets.

Finally, the AMFD data posted online does not include information about the faces' heritage. This information is available upon request to the corresponding author, in order to vet requests and limit the sharing of volunteers' data to

academic researchers only. Furthermore, we believe these faces will be useful for a myriad of social perception experiments in light of the fact that perceivers rarely know targets' racial background when they form initial impressions of them. Nonetheless, the increased protection of heritage information presents a minor inconvenience to researchers and is therefore a limitation of the resource.

In conclusion, we provide the American Multiracial Face Database for free use to all academic researchers. In doing so, we hope to combat existing biases in the Multiracial face perception literature and to contribute to advancing knowledge across the psychological literature of face perception, impression formation, and intergroup relations.

**Acknowledgements** We are grateful to Hyewon Cho, Sua Lee, Karen Martinez, and Cole Russell for their assistance with the standardization and editing of the photographs in the American Multiracial Faces Database and to Daniella Piyavanich, Bowye Tran, and Nizar Helwani for their assistance with data collection. This research was supported by a University of Utah Faculty Research & Creative Grant awarded to the first author.

**Open practices statement** The AMFD faces, data, and codebook for this research are available online (<https://osf.io/qsdrp/>). The AMFD images will be made available to academic researchers for free upon publication of this manuscript. This research was not preregistered because inferential statistics were not appropriate to the research aims.

## References

- Bailey, A. H., LaFrance, M., & Dovidio, J. F. (2019). Is man the measure of all things? A social cognitive account of androcentrism. *Personality and Social Psychology Review*, 23(4), 307–331. <https://doi.org/10.1177/1088868318782848>
- Balas, B., & Pacella, J. (2015). Artificial faces are harder to remember. *Computers in Human Behavior*, 52, 331–337. <https://doi.org/10.1016/j.chb.2015.06.018>
- Balas, B., & Thrash, J. (2019). *Using social face-space to estimate social variables in real and artificial faces*. Unpublished manuscript. <https://doi.org/10.31234/osf.io/btv84>
- Balas, B., Tupa, L., & Pacella, J. (2018). Measuring social variables in real and artificial faces. *Computers in Human Behavior*, 88, 236–243. <https://doi.org/10.1016/j.chb.2018.07.013>
- Brewer, M. B. (1998). A dual process model of impression formation. In T.K. Srull, & R.S. Wyer (Eds.), *Advances in Social Cognition* (pp. 1–36). Lawrence Erlbaum Associates.
- Carpinella, C. M., Chen, J. M., Hamilton, D. L., & Johnson, K. L. (2015). Gendered facial cues influence race categorizations. *Personality and Social Psychology Bulletin*, 41(3), 405–419. <https://doi.org/10.1177/0146167214567153>
- Chen, J. M. (2019). An integrative review of impression formation processes for Multiracial individuals. *Social and Personality Psychology Compass*, 13(1), e12430. <https://doi.org/10.1111/spc3.12430>
- Chen, J.M., Fine, A.D., Norman, J.B., Frick, P.J., & Cauffman, E. (2020). *Out of the picture: Male youths' facial features predict their juvenile justice system processing outcomes*. Manuscript submitted for publication.
- Chen, J. M., & Hamilton, D. L. (2012). Natural ambiguities: Racial categorization of multiracial individuals. *Journal of Experimental*

- Social Psychology*, 48(1), 152–164. <https://doi.org/10.1016/j.jesp.2011.10.005>
- Chen, J. M., Kteily, N. S., & Ho, A. K. (2019). Whose side are you on? Asian Americans' mistrust of Asian–White biracials predicts more exclusion from the ingroup. *Personality and Social Psychology Bulletin*, 45(6), 827–841. <https://doi.org/10.1177/0146167218798032>
- Chen, J. M., Moons, W. G., Gaither, S. E., Hamilton, D. L., & Sherman, J. W. (2014). Motivation to control prejudice predicts categorization of multiracials. *Personality and Social Psychology Bulletin*, 40(5), 590–603. <https://doi.org/10.1177/0146167213520457>
- Chen, J. M., Pauker, K., Gaither, S. E., Hamilton, D. L., & Sherman, J. W. (2018). Black + White = Not White: A minority bias in categorizations of Black–White multiracials. *Journal of Experimental Social Psychology*, 78, 43–54. <https://doi.org/10.1016/j.jesp.2018.05.002>
- Chen, J. M., & Ratliff, K. A. (2015). Implicit attitude generalization from Black to Black–White biracial group members. *Social Psychological and Personality Science*, 6(5), 544–550. <https://doi.org/10.1177/1948550614567686>
- Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist*, 64(3), 170–180. <https://doi.org/10.1037/a0014564>
- Cooley, E., Brown-Iannuzzi, J. L., Brown, C. S., & Polikoff, J. (2018). Black groups accentuate hypodescent by activating threats to the racial hierarchy. *Social Psychological and Personality Science*, 9(4), 411–418. <https://doi.org/10.1177/1948550617708014>
- Dickter, C. L., & Kittel, J. A. (2012). The effect of stereotypical primes on the neural processing of racially ambiguous faces. *Social Neuroscience*, 7(6), 622–631. <https://doi.org/10.1080/17470919.2012.690345>
- Disguised Face Database. (2017). <https://www.cis.upenn.edu/~dfaced/index.html>
- Dunham, Y., & Olson, K. R. (2016). Beyond discrete categories: Studying multiracial, intersex, and transgender children will strengthen basic developmental science. *Journal of Cognition and Development*, 17(4), 642–665. <https://doi.org/10.1080/15248372.2016.1195388>
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science*, 17(5), 383–386. <https://doi.org/10.1111/j.1467-9280.2006.01716.x>
- Evolved Person Perception and Cognition Lab. (n.d.). *Face Stimuli (mostly)*. Retrieved May 2020 from <http://www.epaclab.com/face-stimuli>
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878–902. <https://doi.org/10.1037/0022-3514.82.6.878>
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, 23, 1–74. [https://doi.org/10.1016/S0065-2601\(08\)60317-2](https://doi.org/10.1016/S0065-2601(08)60317-2)
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, 118(2), 247–279. <https://doi.org/10.1037/a0022327>
- Freeman, J. B., Pauker, K., & Sanchez, D. T. (2016). A perceptual pathway to bias: Interracial exposure reduces abrupt shifts in real-time race perception that predict mixed-race bias. *Psychological Science*, 27(4), 502–517. <https://doi.org/10.1177/0956797615627418>
- French, A. R., Franz, T. M., Phelan, L. L., & Blaine, B. E. (2013). Reducing Muslim/Arab stereotypes through evaluative conditioning. *The Journal of Social Psychology*, 153(1), 6–9. <https://doi.org/10.1080/00224545.2012.706242>
- Gaertner, S. L., Dovidio, J. F., & Samuel, G. (2000). Reducing intergroup bias: The common ingroup identity model. Psychology Press.
- Gaither, S. E., Babbitt, L. G., & Sommers, S. R. (2018a). Resolving racial ambiguity in social interactions. *Journal of Experimental Social Psychology*, 76, 259–269. <https://doi.org/10.1016/j.jesp.2018.03.003>
- Gaither, S. E., Chen, J. M., Pauker, K., & Sommers, S. R. (2018b). At face value: Psychological outcomes differ for real vs. computer-generated Multiracial faces. *The Journal of social psychology*, 159(5), 592–610. <https://doi.org/10.1080/00224545.2018.1538929>
- Gaither, S. E., Pauker, K., Slepian, M. L., & Sommers, S. R. (2016). Social belonging motivates categorization of racially ambiguous faces. *Social Cognition*, 34(2), 97–118. <https://doi.org/10.1521/soco.2016.34.2.97>
- Gaither, S. E., Schultz, J. R., Pauker, K., Sommers, S. R., Maddox, K. B., & Ambady, N. (2014). Essentialist thinking predicts decrements in children's memory for racially ambiguous faces. *Developmental Psychology*, 50(2), 482–488. <https://doi.org/10.1037/a0033493>
- Gaither, S. E., Toosi, N. R., Babbitt, L. G., & Sommers, S. R. (2019). Exposure to biracial faces reduces colorblindness. *Personality and Social Psychology Bulletin*, 45(1), 54–66. <https://doi.org/10.1177/0146167218778012>
- Garay, M. M., Meyers, C., Remedios, J. D., & Pauker, K. (2019). Looking like vs. acting like your race: Social activism shapes perceptions of Multiracial individuals. *Self and Identity*, 1–26. <https://doi.org/10.1080/15298868.2019.1659848>
- Ghavami, N., & Peplau, L. A. (2013). An intersectional analysis of gender and ethnic stereotypes: Testing three hypotheses. *Psychology of Women Quarterly*, 37(1), 113–127. <https://doi.org/10.1177/0361684312464203>
- Goff, P. A., Thomas, M. A., & Jackson, M. C. (2008). “Ain't I a woman?": Towards an intersectional approach to person perception and group-based harms. *Sex Roles*, 59(5–6), 392–403. <https://doi.org/10.1007/s11199-008-9505-4>
- Grgic, M., & Delac, K. (2019). *Face Recognition Databases*. <https://www.face-rec.org/databases/>
- Gunaydin, G., Selcuk, E., & Zayas, V. (2017). Impressions based on a portrait predict, 1-month later, impressions following a live interaction. *Social Psychological and Personality Science*, 8(1), 36–44. <https://doi.org/10.1177/1948550616662123>
- Halberstadt, J., Sherman, S. J., & Sherman, J. W. (2011). Why Barack Obama is Black: A cognitive account of hypodescent. *Psychological Science*, 22(1), 29–33. <https://doi.org/10.1177/0956797610390383>
- Halberstadt, J., & Winkielman, P. (2014). Easy on the eyes, or hard to categorize: Classification difficulty decreases the appeal of facial blends. *Journal of Experimental Social Psychology*, 50, 175–183. <https://doi.org/10.1016/j.jesp.2013.08.004>
- Helman, E., Sutherland, C. A., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Journal of Personality and Social Psychology*, 113(4), 513–529. <https://doi.org/10.1037/pspa0000090>
- Helman, E., Xie, S. Y., Ofofu, E. K., & Nespoli, G. A. (2018). *Assessing the point at which averages are stable: A tool illustrated in the context of person perception*. Unpublished manuscript. 10.31234/osf.io/2n6jq
- Herman, M. R. (2010). Do you see what I am? How observers' backgrounds affect their perceptions of Multiracial faces. *Social Psychology Quarterly*, 73(1), 58–78. <https://doi.org/10.1177/0190272510361436>
- Hinzman, L., & Kelly, S. D. (2013). Effects of emotional body language on rapid out-group judgments. *Journal of Experimental Social Psychology*, 49(1), 152–155. <https://doi.org/10.1016/j.jesp.2012.07.010>
- Ho, A. K., Kteily, N. S., & Chen, J. M. (2020). Introducing the sociopolitical motive × intergroup threat model to understand how monoracial perceivers' sociopolitical motives influence their



- categorization of multiracial people. *Personality and Social Psychology Review*, 1-27. <https://doi.org/10.1177/1088868320917051>.
- Ho, A. K., Roberts, S. O., & Gelman, S. A. (2015). Essentialism and racial bias jointly contribute to the categorization of multiracial individuals. *Psychological Science*, 26(10), 1639-1645. <https://doi.org/10.1177/0956797615596436>
- Ho, A. K., Sidanius, J., Levin, D. T., & Banaji, M. R. (2011). Evidence for hypodescent and racial hierarchy in the categorization and perception of biracial individuals. *Journal of Personality and Social Psychology*, 100(3), 492. <https://doi.org/10.1037/a0021562>
- Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science*, 15(5), 342-345. <https://doi.org/10.1111/j.0956-7976.2004.00680.x>
- Hutchings, P. B., & Haddock, G. (2008). Look Black in anger: The role of implicit prejudice in the categorization and perceived emotional intensity of racially ambiguous faces. *Journal of Experimental Social Psychology*, 44(5), 1418-1420. <https://doi.org/10.1016/j.jesp.2008.05.002>
- Ito, T. A., Willadsen-Jensen, E. C., Kaye, J. T., & Park, B. (2011). Contextual variation in automatic evaluative bias to racially-ambiguous faces. *Journal of Experimental Social Psychology*, 47(4), 818-823. <https://doi.org/10.1016/j.jesp.2011.02.016>
- Jaeger, B., Evans, A. M., Stel, M., & van Beest, I. (2019). Explaining the persistent influence of facial cues in social decision-making. *Journal of Experimental Psychology: General*, 148(6), 1008–1021. <https://doi.org/10.1037/xge0000591>
- Jimeno-Ingrum, D., Berdahl, J. L., & Lucero-Wagoner, B. (2009). Stereotypes of Latinos and Whites: Do they guide evaluations in diverse work groups? *Cultural Diversity and Ethnic Minority Psychology*, 15(2), 158–164. <https://doi.org/10.1037/a0015508>
- Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, 102(1), 116-131. <https://doi.org/10.1037/a0025335>
- Jones, B., DeBruine, L., Flake, J., Aczel, B., Adamkovic, M., Alaei, R., ... Chartier, C. R. (2019). To Which World Regions Does the Valence-Dominance Model of Social Perception Apply? *Nature Human Behaviour*.
- Kang, S. K., Plaks, J. E., & Remedios, J. D. (2015). Folk beliefs about genetic variation predict avoidance of biracial individuals. *Frontiers in Psychology*, 6(357), 1-11. <https://doi.org/10.3389/fpsyg.2015.00357>
- Kenny, D. A., & West, T. V. (2010). Similarity and agreement in self-and other perception: A meta-analysis. *Personality and Social Psychology Review*, 14(2), 196-213. <https://doi.org/10.1177/1088868309353414>
- Kinzler, K. D., & Dautel, J. B. (2012). Children's essentialist reasoning about language and race. *Developmental Science*, 15(1), 131-138. <https://doi.org/10.1111/j.1467-7687.2011.01101.x>
- Krosch, A. R., Berntsen, L., Amodio, D. M., Jost, J. T., & Van Bavel, J. J. (2013). On the ideology of hypodescent: Political conservatism predicts categorization of racially ambiguous faces as Black. *Journal of Experimental Social Psychology*, 49(6), 1196-1203. <https://doi.org/10.1016/j.jesp.2013.05.009>
- Kteily, N., Cotterill, S., Sidanius, J., Sheehy-Skeffington, J., & Bergh, R. (2014). "Not one of us" predictors and consequences of denying ingroup characteristics to ambiguous targets. *Personality and Social Psychology Bulletin*, 40(10), 1231-1247. <https://doi.org/10.1177/0146167214539708>
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, 24(8), 1377-1388. <https://doi.org/10.1080/02699930903485076>
- Livingston, G. (2017). *The rise of Multiracial and multiethnic babies in the U.S.* Pew Research Center. <https://www.pewresearch.org/fact-tank/2017/06/06/the-rise-of-Multiracial-and-multiethnic-babies-in-the-u-s/>
- LoBue, V., & Thrasher, C. (2015). The Child Affective Facial Expression (CAFE) set: Validity and reliability from untrained adults. *Frontiers in Psychology*, 5(1532), 1-8. <https://doi.org/10.3389/fpsyg.2014.01532>
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47(4), 1122-1135. <https://doi.org/10.3758/s13428-014-0532-5>
- Ma, D. S., Koltai, K., McManus, R. M., Bernhardt, A., Correll, J., & Wittenbrink, B. (2018). Race signaling features: Identifying markers of racial prototypicality among Asians, Blacks, Latinos, and Whites. *Social Cognition*, 36(6), 603-625. <https://doi.org/10.1521/soco.2018.36.6.603>
- MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: the ambiguous race face effect. *Psychology, Public Policy, and Law*, 7(1), 98-118. <https://doi.org/10.1037/1076-8971.7.1.98>
- MacLin, O. H., & Malpass, R. S. (2003). The ambiguous-race face illusion. *Perception*, 32(2), 249-252. <https://doi.org/10.1068/p5046>
- Mattarozzi, K., Colonnello, V., De Gioia, F., & Todorov, A. (2017). I care, even after the first impression: Facial appearance-based evaluations in healthcare context. *Social Science & Medicine*, 182, 68-72. <https://doi.org/10.1016/j.socscimed.2017.04.011>
- Miller, S. L., Maner, J. K., & Becker, D. V. (2010). Self-protective biases in group categorization: Threat cues shape the psychological boundary between "us" and "them". *Journal of Personality and Social Psychology*, 99(1), 62-77. <https://doi.org/10.1037/a0018086>
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, 36(4), 630-633. <https://doi.org/10.3758/BF03206543>
- Newton, V. A., Dickter, C. L., & Gyurovski, I. I. (2011). The effects of stereotypical cues on the social categorization and judgment of ambiguous-race targets. *Journal of Interpersonal Relations, Intergroup Relations and Identity*, 4, 31-45.
- Nicolas, G., Skinner, A. L., & Dickter, C. L. (2019). Other Than the Sum: Hispanic and Middle Eastern Categorizations of Black-White Mixed-Race Faces. *Social Psychological and Personality Science*, 10(4), 532-541. <https://doi.org/10.1177/1948550618769591>
- Norman, J. B., & Chen, J. M. (2019). I am Multiracial: Predictors of Multiracial identification strength among mixed ancestry individuals. *Self and Identity*, 1-20. <https://doi.org/10.1080/15298868.2019.1635522>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087-11092. <https://doi.org/10.1073/pnas.0805664105>
- Parker, K., Horowitz, J. M., Morin, R., & Lopez, M. H. (2015). *Multiracial in America: Proud, diverse and growing in Numbers.* Pew Research Center. <https://www.pewsocialtrends.org/2015/06/11/Multiracial-in-america/>
- Pauker, K., & Ambady, N. (2009). Multiracial faces: How categorization affects memory at the boundaries of race. *Journal of Social Issues*, 65(1), 69-86. <https://doi.org/10.1111/j.1540-4560.2008.01588.x>
- Pauker, K., Ambady, N., & Freeman, J. B. (2013a). The power of identity to motivate face memory in biracial individuals. *Social Cognition*, 31(6), 780-791. <https://doi.org/10.1521/soco.2013.31.6.780>
- Pauker, K., Carpinella, C. M., Lick, D. J., Sanchez, D. T., & Johnson, K. L. (2018a). Malleability in biracial categorizations: The impact of geographic context and targets' racial heritage. *Social Cognition*, 36(5), 461-480. <https://doi.org/10.1521/soco.2018.36.5.461>
- Pauker, K., Johnson, K. L., & Ambady, N. (2013b). *Race in the mirror: Situational cues and essentialist theories influence biracial individuals' self-perception.* Manuscript in preparation.

- Pauker, K., Meyers, C., Sanchez, D. T., Gaither, S. E., & Young, D. M. (2018b). A review of multiracial malleability: Identity, categorization, and shifting racial attitudes. *Social and Personality Psychology Compass*, 12(6), e12392. <https://doi.org/10.1111/spc3.12392>
- Pauker, K., Weisbuch, M., Ambady, N., Sommers, S. R., Adams Jr, R. B., & Ivcevic, Z. (2009). Not so black and white: memory for ambiguous group members. *Journal of Personality and Social Psychology*, 96(4), 795–810. <https://doi.org/10.1037/a0013265>
- Peery, D., & Bodenhausen, G. V. (2008). Black+ White= Black: Hypodescent in reflexive categorization of racially ambiguous faces. *Psychological Science*, 19(10), 973–977. <https://doi.org/10.1111/j.1467-9280.2008.02185.x>
- Plaks, J. E., Malahy, L. W., Sedlins, M., & Shoda, Y. (2012). Folk beliefs about human genetic variation predict discrete versus continuous racial categorization and evaluative bias. *Social Psychological and Personality Science*, 3(1), 31–39. <https://doi.org/10.1177/1948550611408118>
- Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles*, 59, 377–391. <https://doi.org/10.1007/s11199-008-9424-4>
- Roberts, S. O., & Gelman, S. A. (2015). Do children see in Black and White? Children's and adults' categorizations of multiracial individuals. *Child Development*, 86(6), 1830–1847. <https://doi.org/10.1111/cdev.12410>
- Roberts, S. O., & Gelman, S. A. (2017). Multiracial children's and adults' categorizations of multiracial individuals. *Journal of Cognition and Development*, 18(1), 1–15. <https://doi.org/10.1080/15248372.2015.1086772>
- Roberts, S. O., Leonard, K. C., Ho, A. K., & Gelman, S. A. (2017a). Does this Smile Make me Look White? Exploring the Effects of Emotional Expressions on the Categorization of Multiracial Children. *Journal of Cognition and Culture*, 17(3–4), 218–231. <https://doi.org/10.1163/15685373-12340005>
- Roberts, S. O., Williams, A. D., & Gelman, S. A. (2017b). Children's and adults' predictions of Black, White, and Multiracial friendship patterns. *Journal of Cognition and Development*, 18(2), 189–208. <https://doi.org/10.1080/15248372.2016.1262374>
- Rodeheffer, C. D., Hill, S. E., & Lord, C. G. (2012). Does this recession make me look black? The effect of resource scarcity on the categorization of biracial faces. *Psychological Science*, 23(12), 1476–1478. <https://doi.org/10.1177/0956797612450892>
- Sanchez, D. T., & Chavez, G. (2010). Are you minority enough? Language ability affects targets' and perceivers' assessments of a candidate's appropriateness for affirmative action. *Basic and Applied Social Psychology*, 32(1), 99–107. <https://doi.org/10.1080/01973530903435896>
- Sanchez, D. T., Chavez, G., Good, J. J., & Wilton, L. S. (2012). The language of acceptance: Spanish proficiency and perceived intragroup rejection among Latinos. *Journal of Cross-Cultural Psychology*, 43(6), 1019–1033. <https://doi.org/10.1177/0022022111416979>
- Sesko, A. K., & Biernat, M. (2010). Prototypes of race and gender: The invisibility of Black women. *Journal of Experimental Social Psychology*, 46(2), 356–360. <https://doi.org/10.1016/j.jesp.2009.10.016>
- Shutts, K., & Kinzler, K. D. (2007). An ambiguous-race illusion in children's face memory. *Psychological Science*, 18(9), 763–767. <https://doi.org/10.1111/j.1467-9280.2007.01975.x>
- Skinner, A. L., & Nicolas, G. (2015). Looking Black or looking back? Using phenotype and ancestry to make racial categorizations. *Journal of Experimental Social Psychology*, 57, 55–63. <https://doi.org/10.1016/j.jesp.2014.11.011>
- Small, P. A., & Major, B. (2019). Crossing the racial line: The fluidity vs. fixedness of racial identity. *Self and Identity*, 1–26. <https://doi.org/10.1080/15298868.2019.1662839>
- Stolier, R. M. (n.d.). *FSTC: Face stimulus and tool collection*. Retrieved May 2020, from <https://rystoli.github.io/FSTC.html#stim>
- Stroessner, S. J. (1996). Social categorization by race or sex: Effects of perceived non-normalcy on response times. *Social Cognition*, 14(3), 247–276. <https://doi.org/10.1521/soco.1996.14.3.247>
- Strohinger, N., Gray, K., Chituc, V., Heffner, J., Schein, C., & Heagins, T. B. (2016). The MR2: A multi-racial, mega-resolution database of facial stimuli. *Behavior Research Methods*, 48(3), 1197–1204. <https://doi.org/10.3758/s13428-015-0641-9>
- Tajfel, H., & Turner, J. (1979). An integrative theory of group conflict. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33–47). Brooks/Cole.
- Thorntson, C. A., Pazda, A. D., Young, S. G., & Slepian, M. L. (2019). Incidental Cues to Threat and Racial Categorization. *Social Cognition*, 37(4), 389–404. <https://doi.org/10.1521/soco.2019.37.4.389>
- Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, 12(12), 455–460. <https://doi.org/10.1016/j.tics.2008.10.001>
- Willadsen-Jensen, E., & Ito, T. A. (2015). The effect of context on responses to racially ambiguous faces: changes in perception and evaluation. *Social Cognitive and Affective Neuroscience*, 10(7), 885–892. <https://doi.org/10.1093/scan/nsu134>
- Willadsen-Jensen, E. C., & Ito, T. A. (2006). Ambiguity and the timecourse of racial perception. *Social Cognition*, 24(5), 580–606. <https://doi.org/10.1521/soco.2006.24.5.580>
- Willadsen-Jensen, E. C., & Ito, T. A. (2008). A foot in both worlds: Asian Americans' perceptions of Asian, White, and racially ambiguous faces. *Group Processes & Intergroup Relations*, 11(2), 182–200. <https://doi.org/10.1177/1368430207088037>
- Williams, M. J., George-Jones, J., & Hebl, M. (2019). The face of STEM: Racial phenotypic stereotypicality predicts STEM persistence by—and ability attributions about—students of color. *Journal of Personality and Social Psychology*, 116(3), 416–443. <https://doi.org/10.1037/pspi0000153>
- Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, 26(8), 1325–1331. <https://doi.org/10.1177/0956797615590992>
- Wilson, J. P., & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes: The role of facial trustworthiness. *Social Psychological and Personality Science*, 7(4), 331–338. <https://doi.org/10.1177/1948550615624142>
- Wilton, L., Sanchez, D. T., & Giammo, L. (2014). Seeing Similarity or Distance?: Racial Identification Moderates Intergroup Perception after Biracial Exposure. *Social Psychology*, 45, 127–134. <https://doi.org/10.1027/1864-9335/a000168>
- Wilton, L. S., Rattan, A., & Sanchez, D. T. (2018). White's perceptions of biracial individuals' race shift when biracials speak out against bias. *Social Psychological and Personality Science*, 9(8), 953–961. <https://doi.org/10.1177/1948550617731497>
- Wilton, L. S., Sanchez, D. T., & Chavez, G. F. (2013). Speaking the language of diversity: Spanish fluency, White ancestry, and skin color in the distribution of diversity awards to Latinos. *Basic and Applied Social Psychology*, 35(4), 346–359. <https://doi.org/10.1080/01973533.2013.803969>
- Xie, S. Y., Flake, J. K., & Hehman, E. (2019). Perceiver and target characteristics contribute to impression formation differently across race and gender. *Journal of Personality and Social Psychology*, 117(2), 364–385. <https://doi.org/10.1037/pspi0000160>

- Young, D. M., Sanchez, D. T., & Wilton, L. S. (2013). At the crossroads of race: Racial ambiguity and biracial identification influence psychological essentialist thinking. *Cultural Diversity and Ethnic Minority Psychology, 19*(4), 461. <https://doi.org/10.1037/a0032565>
- Young, D. M., Sanchez, D. T., & Wilton, L. S. (2016). Too rich for diversity: Socioeconomic status influences multifaceted person perception of Latino targets. *Analyses of Social Issues and Public Policy, 16*(1), 392–416. <https://doi.org/10.1111/asap.12104>
- Young, D. M., Sanchez, D. T., & Wilton, L. S. (2017). Biracial perception in black and white: How Black and White perceivers respond to phenotype and racial identity cues. *Cultural Diversity and Ethnic Minority Psychology, 23*(1), 154–164. <https://doi.org/10.1037/cdp0000103>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.