

## Hearing in Color: How Expectations Distort Perception of Skin Tone

Ulrik Lyngs<sup>1</sup>, Emma Cohen<sup>2,3</sup>, Wallisen Tadashi Hattori<sup>4</sup>, Martha Newson<sup>2</sup>, Daniel T. Levin<sup>5</sup>

### Institutions

<sup>1</sup> Aarhus University

<sup>2</sup> University of Oxford, <sup>3</sup>Wadham College

<sup>4</sup> Federal University of Uberlândia

<sup>5</sup> Vanderbilt University

### Author Note

Ulrik Lyngs, School of Culture and Society, Aarhus University; Emma Cohen, Wadham College, Oxford & Institute of Cognitive and Evolutionary Anthropology, University of Oxford; Martha Newson, Institute of Cognitive and Evolutionary Anthropology, University of Oxford; Wallisen Tadashi Hattori, Department of Public Health, Medical School, Federal University of Uberlândia; Daniel Levin, Department of Psychology and Human Development, Vanderbilt University.

This research was supported by funding from Peter Lienhardt Memorial Fund and Philip Bagby Fund to U.L., the Brazilian Coordination for the Improvement of Higher Education (CAPES PNPD Process N° 2925/2010) to W.T.H., and a British Academy Fellowship to E.C.

Correspondence concerning this article should be addressed to Ulrik Lyngs at [ulrik.lyngs@post.au.dk](mailto:ulrik.lyngs@post.au.dk)

### Abstract

Previous research has found that the perceived brightness of a face can be distorted by the social category of race. Thus, Levin & Banaji (2006) found in a US sample that faces of identical brightness were perceived to be lighter if they had stereotypical White American features than if they had Black American features. Here, we present two experiments conducted in Natal, Brazil, that extend this line of research. Experiment 1 tested if the brightness distortion effect would generalize to a Brazilian population. Experiment 2 tested if speech accent would have a similar effect on brightness perception. In Experiment 1, we found that the brightness distortion effect clearly replicated in the Brazilian sample: faces with Black racial features were perceived to be darker than faces with White racial features, even though their objective brightness was identical. In Experiment 2, we found that speech accent influenced brightness perception in a similar manner: faces were perceived to be darker when paired with an accent associated with low socio-economic status than when they were paired with an accent associated with high socio-economic status. Whereas racial concepts in Brazil are often claimed to be much more fluid compared to the US, our findings suggest that the populations are quite similar with respect to associations between facial features and skin tone. Our findings also demonstrate speech accent as an additional source of category-information that perceptual cognition can take into account when modelling the world.

## **Hearing in Color: How Expectations Distort Perception of Skin Tone**

A long tradition of research on human perception has established that basic perceptual processes are affected by knowledge about the world. For example, knowledge about the human body affects how we perceive the motion of limbs (Shiffrar & Freyd, 1993), the familiarity of contours can determine whether we perceive them as part of a figure or as part of the background (Peterson, 1994), and the perception of an object's color is highly dependent on its context (Lotto & Purves, 2002). Extending this tradition, Levin & Banaji (2006) demonstrated that the social category of race can distort the perceived lightness of faces. They showed that faces of identical on-screen reflectance are perceived to be lighter if the faces have stereotypical White American (henceforth simply 'White') facial features than if they have African American ('Black') features. Here, we test whether this effect is moderated by two important forms of cultural context: the conceptual structure of race categories and the relationship between skin color and characteristic speech. Specifically, whereas Levin & Banaji conducted their research with American students from Kent State University, we investigated whether the findings would replicate in Brazil, where racial categories are arguably more heterogeneous than in the US (e.g., Fish 2005). In addition, we investigated how perception of facial skin tone is influenced by accent cues, inspired by studies of how speech accent influences social categorization and cognition (e.g. Rakić et al., 2011; Dehghani et al., 2014).

### **Contextual influences on brightness perception**

Starting with Helmholtz's (1866) work on lightness constancy, and continuing with the anchoring theory of lightness perception by Gilchrist and colleagues (1999; see Adelson 2000 for a review), one central research problem has been to understand how our perception

of the relative reflectance (i.e. lightness) of different surfaces in our surroundings remains stable across great variation in ambient light conditions (Adelson, 2000). Here, a number of well-known illustrations have shown that perception of lightness is not just an objective readout of a given object's luminance, but is contingent on a number of contextual features. One line of explanation focuses on early processes in the visual system, such as the patterns of excitation and inhibition in retinal ganglion cell receptive fields (Cornsweet, 1970). Another focuses on higher-level inferences that include a wider range of more sophisticated perceptual factors, such as shadows and transmittal cues (Purves, Williams, Nundy, & Lotto, 2004).

Until recently, little research had tested how an object's category affects lightness perception (cf. Levin & Banaji, 2006). A first set of studies by MacLin & Malpass (2001; 2003) included an exploration of whether social classification affects perceived lightness. This study demonstrated that different characteristic hair styles led participants to classify otherwise ambiguous faces as "Black" or "Hispanic", and that these classifications were associated with judgments that the Black faces were darker. However, these studies relied upon rating scale judgments of how "light" or "dark" the faces were, which may have reflected a nonperceptual judgment rather than a true brightness effect. Therefore, Levin & Banaji (2006) used a more traditional perceptual adjustment paradigm to test the same hypothesis. In Levin & Banaji's experiments, participants adjusted the lightness of a visual sample to match a face they classified as Black or White. Specifically, participants saw a face on a computer screen with an adjustable stimulus (either a face or a grey square) to the right, and used the up- and down-arrow keys to make the stimulus lighter or darker until they judged it to be similar to the skin tone of the face. The faces were identical in objective reflectance – only features diagnostic of race differed (e.g. shape and size of mouth, nose,

jaw).<sup>1</sup> The experiments demonstrated a robust lightness distortion, with participants consistently choosing darker samples for Black faces than for White faces – an effect which was significant for both Black and White participants when analyzed independently.

So far, this distortion effect has only been demonstrated in the US. However, since the effect is assumed to be based on learned associations between racial categories and average skin color (Levin & Banaji, 2006), it is an open question whether the effect extends to populations with different ethnic compositions. Hence, the first question that we address in the present paper is whether the effect generalizes to a population in Brazil. Brazil has long been claimed to have more ambiguous, fluid, and ‘fuzzy’ racial categories than the US (Harris et al. 1993; 1995; Byrne et al. 1995; however see Gil-White, 2001 or Jones, 2009 for a critical discussion). Therefore, to the extent that Brazilians have a more variegated, less binary perception of racial categories than Americans, and base provisional categorization on a wider range of traits, it may be hypothesized that significant and strong distortion of perceived skin color based on facial morphology alone is less likely among this population (cf. Telles & Sue, 2009; Telles, 2004). Alternatively, recent policies promoting racial equality in Brazil, including, for example, the introduction of a quota system in tertiary education and public office, may have served to raise consciousness of race-related categories and promote starker divides in social perception. For these reasons, Brazil offers a relevant and interesting opportunity to explore how cultural context influences the distortion effect of race category on skin color perception.

Although essentially all previous lightness perception research has explored visual inputs to brightness, the finding that object categories can impact brightness suggests that a broad range of properties associated with those categories may affect basic visual judgments.

---

<sup>1</sup> In one of their experiments, participants even rated lightness of the exact same, racially ambiguous, face differently depending on whether the face was labelled ‘Black’ or ‘White’.

This is particularly true of faces, which are sources of spoken language, where speech patterns in turn are powerful cues to ethnicity and social position (Bayard, Weatherall, Gallois, & Pittam, 2001; Foon, 2001). An extensive set of research findings demonstrates that face perception and speech perception are governed by distinct, but coordinated neural systems (for review see Haxby, Hoffman, & Gobbini, 2000). Thus, another key aspect we explore is the influence that variations in characteristic speech patterns may have upon lightness perception of faces. A number of previous studies have explored how accent is used for social categorization, and how it can influence cognition and behavior. Kinzler, Shutts, DeJesus, & Spelke (2009) found that 5-year-old children preferred native speakers of their own language as friends over those who spoke with a foreign accent. This held true even when accent preferences were pitted against skin color preferences, leading the authors to conclude that accent ‘trumps’ race in children's early social preferences. Similar accent-based preferences for local over non-local speakers (rather than foreign vs. native) were found among Brazilian children raised in multi-accent environments, but not among children in linguistically homogeneous environments (Cohen & Haun, 2013). Using a ‘Who Said What’ paradigm, Rakić et al. (2011) found that people base ethnic categorization on accent, even when visual cues to ethnicity cross-cut accent groupings. Other studies using this paradigm have found similar results, suggesting that accent is a dedicated dimension of social categorization (Pietraszewski & Schwartz, 2014a, 2014b; see also Cohen, 2012). A recent study by Dehghani et al. (2014) further found that bicultural (i.e. individuals who have internalized two distinct cultural identities) Americans could be primed to change their interpretive frame and social decision-making depending on whether a situation was verbally presented in English with an American accent or with the accent of the participants’ alternative cultural identity (in this study Chinese, Iranian or Mexican).

These studies, building on the extensive literature in sociolinguistics more generally (e.g., Giles, 2012), suggest that a speaker's accent carries important information that biases listeners' social categorizations, interpersonal evaluations and behavior. A central point for the present paper is that accent is often a strong cue to socioeconomic status (cf. Bayard, Weatherall, Gallois, & Pittam, 2001; Foon, 2001; Fuertes, Potere, & Ramirez, 2002). In Brazil, socio-economic status is in turn highly correlated with skin color (Schwartzman, 2007), as reflected in the common idea, reported in the sociological literature on Brazil, that "money whitens" (cf. Degler, 1971). Hence, Brazil is a particularly suitable context in which to explore how both race and accent influence lightness perception.

### **Summary and hypotheses**

We carried out two experiments in Natal, Brazil. Natal is the capital city of the northern state of Rio Grande do Norte with nationally representative race-based disparities in SES. For example, the 2010 census reported that, among those over the age of 15, 13.2% of the Black population was illiterate as compared with 5.6% of the White population (IBGE, 2010). The average monthly salary of self-identified Black respondents was around half that of White respondents (995 and 1845 Brazilian Reals, respectively). It therefore provided an appropriate context in which to locate our investigation the effects of class and race on face lightness judgments.

First, we re-ran Levin & Banaji (2006)'s work. As described in the preceding section, it was important to investigate the brightness distortion effect cross-culturally, as previous anthropological research on race categories in Brazil (Byrne et al., 1995; Harris et al., 1993, 1995) suggests that the learned associations underlying the effect might differ in this target population, relative to the US. In addition to testing whether the effect generalizes across cultures, we also tested whether the effect generalizes across stimuli. The original Levin and Banaji experiments relied upon a single pair of highly controlled average faces that were

matched between the races for mean luminance and contrast. For this experiment, we generated another pair of Black and White average faces based on a different set of original faces. A demonstration of the lightness effect in these faces would not only confirm that the original lightness effect is not unique to a particular stimulus set, but would also support our goals in Experiment 2 which required additional validated stimuli.

A secondary goal of Experiment 1 was to assess the degree to which the brightness effect would also generalize to non-average faces. To this end, we created blend faces that were distorted to look more like faces of the other race. Thus, a further version of each Black average face was created by generating new blends that represented a location on a Black-White continuum that was 35% of the distance from the Black face to the White face, and the converse was done for each White face. These faces will be referred to as Black blends (i.e. more Black than White) and White blends (i.e. more White than Black). They represent faces for which visible racial features are less pronounced, but remain easily identifiable to observers as White or Black. The original undistorted average faces will be referred to as prototypes.

In Experiment 2, participants judged the lightness of faces that were paired with sentences spoken in distinct accents. We expected to observe diminished brightness distortion for face-accent pairings where visual features and speech carried opposite associations about likely skin tone. As a sub-goal, we also investigated whether participants' recall of voice-face pairings would reveal how they encoded voice-race relations.

## Method

**Participants.** 44 participants (18 female) were recruited on campus at Universidade Federal do Rio Grande do Norte in Natal, Brazil. 22 indicated that they were white (*branca*), 20 that they were brown (*parda*), and 2 that they were black (*preta*). Mean age was 23.4 years (range 18-33 years).



**Apparatus.** Half of the participants were exposed to the experimental stimuli on a 2013 15” Macbook Pro with ‘Retina’ display, the other half on a 20-inch iMac. Both screens were set at a resolution of 1680 by 1050 pixels, and at the highest brightness setting. For the accent-association task at the end of the experiment, sound recordings were played to the participants from a 2008 Macbook Air using earphones.

**Stimuli.** Subjects rated two different pairs of Black and White faces that were matched between the races for mean luminance and contrast<sup>2</sup>. Set 1 was adopted from Levin & Banaji (2006), and included two prototype faces, one White and one Black. The prototype faces were originally created by blending a set of 16 faces for each race (for details, see Levin & Banaji, 2006, p. 503). In addition to the prototypes, we used the program MorphX to create a pair of blended faces that each represented a point on a form and texture map continuum between the prototypes that was 35% of the distance from one face to the other. Thus, we created a variant of the Black prototype that was 35% of the distance from that face to the White face and a White face that was 35% of the distance from that face to the Black face. These faces remain identifiable as members of each race, although the race-specifying information is more ambiguous (see Fig. 1). As in Levin and Banaji (2006) we created 4 brightness variants of each face, two that were lighter by 10 and 20 levels (relative to the total range of 256 grey-scale levels) and two that were darker by 10 and 20 levels. In addition, we created another set of faces using the same process. The second set of prototypes were

---

<sup>2</sup> We note that using greyscale images that lack hue information may not perfectly reflect known differences between different face categories, and have done so because the process of equalizing luminance in color images is complex, and because we follow an extensive tradition of research exploring luminance in isolation. However, research demonstrating the perceptual impact of knowledge about object categories leads us to believe that our effects would be very similar with color images (see for example, Mitterer & de Ruiter, 2008).

combinations of 8 Black and 8 White head shots of bald men taken from the book *Heads* (Kayser, 1997). These photographs were chosen because they are all full-frontal images with neutral expressions.

For both sets of faces luminance was matched between the races by matching the mean and standard deviation of 8-bit brightness values, as measured in Adobe Photoshop CS6. For the set 1 faces, new versions of the Levin and Banaji (2006) faces were rematched (the originals had been stored in the now-obsolete .pict format). The White prototype had a mean grey level of 120.43 (in an 8-bit, 256-level range;  $SD = 39.24$ ) and the Black prototype a mean level of 120.73 ( $SD = 39.20$ ). The White blend had a mean grey level of 120.71 ( $SD = 39.20$ ), and the Black blend had a mean grey level of 120.52 ( $SD = 39.15$ ). For set 2, the White prototype had a mean level of 134.53 ( $SD = 39.37$ ) and the Black prototype had a mean level of 134.53 ( $SD = 39.37$ ). The White blend had a mean grey level of 135.29 ( $SD = 38.40$ ), and the Black blend had a mean grey level of 135.43 ( $SD = 38.71$ ). Lighter and darker variants of these faces were created by increasing and decreasing mean grey scale values by 20 levels in 10-level increments. This was done by adding a constant of 10 levels to all greyscale values (using the legacy brightness adjustment in Photoshop CS6). Adjustable sample regions (that is, the grey squares that participants matched to the faces in the experimental task) were uniform grey patches that ranged in 17 5-level steps from a grey level of 101 at the darkest to 181 at the lightest.

The experiment was programmed in PsychoPy, stimulus presentation software based on Python (Peirce, 2007).

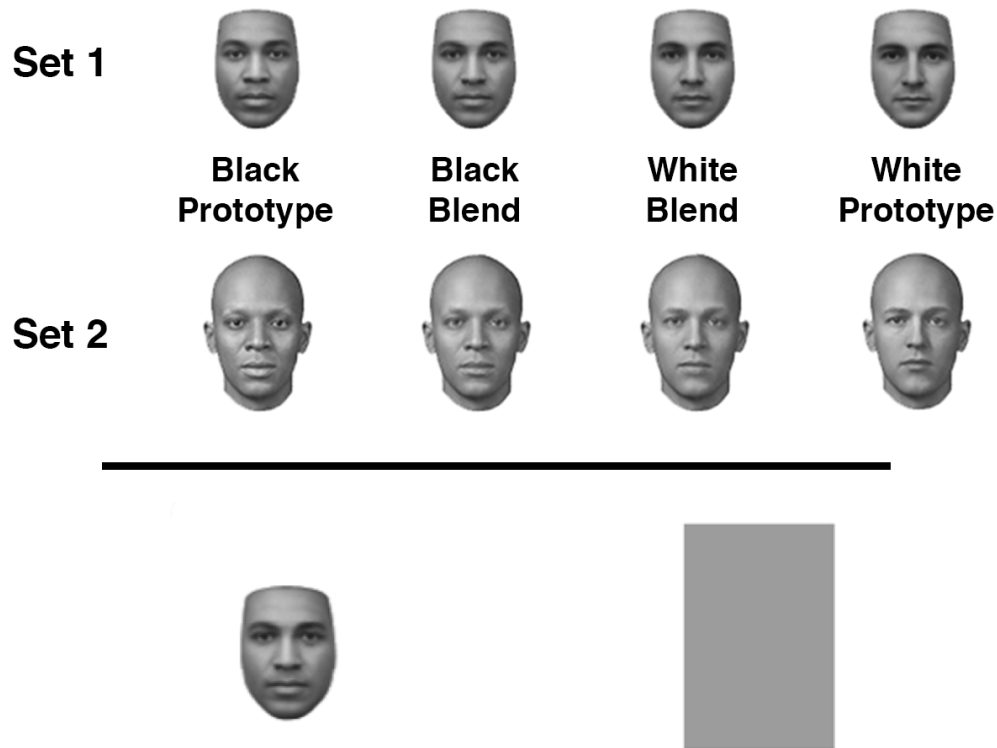


Figure 1. Top: Set 1 and set 2 prototype faces and blends. Bottom: Black average face adjacent to adjustable sample.

**Procedure.** Participants were recruited for a study on ‘visual perception of faces’. Participants were tested in small groups, with two subjects completing the experiment simultaneously on separate computers. First, a research assistant recorded the age and sex of the participant on the computer. Participants then read instructions on the screen and began the experimental task when they were ready (for the instructions given see Appendix A).

For each trial, a face appeared in the middle of the screen adjacent to an adjustable grey square to the right. Participants were instructed to adjust the color of the square using the up- and down- arrows on the keyboard until they judged it to be as similar as possible to the overall skin color of the face, and then to press the spacebar to move on to the next trial. The experiment included 80 trials, 40 for each set of faces, with the two sets presented in

separate blocks. Within each block, the order of trials was randomized, and the order of the blocks was counterbalanced between participants.

After the experimental tasks, participants completed a brief measure of racial attitudes, using a variant of the ‘feelings thermometer’. This measure, in which participants rate their attitudes to various groups, adapted the scale used by Levin & Banaji (2006) to the local context. Alongside "filler groups" (e.g., political parties, football supporters) were racially defined categories (Black/Afro-Brazilian, White, Brown (*Pardos/morenas*)). Participants rated their attitudes by using a ‘thermometer’, entering a number between 0 and 100, where 0 corresponded to the anchor point *very coolly* (*muito frio*) and 100 to the anchor point *very warmly* (*muito caloroso*).

Participants then completed additional demographic information and an ‘accent association task’ unrelated to experiment 1 (testing the accents to be used in experiment 2). In this task, the participant listened to recordings of four different voices and filled in a questionnaire about which color and income level they associated with the voices.

## Results

The distortion in lightness perception reported by Levin & Banaji (2006) was replicated overall and in each set of faces (see Figure 2). Participants consistently chose darker samples for the Black reference faces (mean of 8.07,  $SE = 0.34$ ; 0 is darkest possible option, 16 lightest possible) than for the White faces ( $M = 10.13$ ,  $SE = 0.27$ ),  $t(43) = -6.0$ ,  $p < .001$ ,  $d = 1.00$ . 43 of 44 participants (98%) chose darker samples for Black faces than for White faces. The effect was consistent across both sets of stimulus faces. For Set 1 (the faces adapted from Levin & Banaji 2006), participants on average chose a sample 2.50 levels darker for the Black ( $M = 7.77$ ,  $SE = 0.37$ ) than for the White faces ( $M = 10.26$ ,  $SE = 0.31$ ),  $t(43) = -7.02$ ,  $p < .001$ ,  $d = 1.11$ . For Set 2, participants on average chose a sample 1.63 levels darker for the Black ( $M = 8.37$ ,  $SE = 0.38$ ) than for the White faces ( $M = 10.00$ ,  $SE = 0.29$ ),

$t(43) = -4.48, p < .001, d = 0.72$ . Participants' self-identified color did not influence the size of the distortion effect (comparing *branca* and *parda/preta*),  $t(42) = -0.41, p = .687$ . The effect was more consistent across participants in Set 1 (43 out of 44 participants chose darker samples for Black faces than White faces) than in Set 2 (33 out of 44;  $p = .004$ , Fischer's exact test.).

The lightness effect was significant both for undistorted prototype stimuli and for the blends (see Figure 2). For prototype stimuli, participants chose samples 2.70 levels darker for Black faces ( $t(43) = 6.48, p < .001, d = 1.22$ ), and for blend faces they chose samples 1.42 levels darker for Black faces ( $t(43) = 4.93, p < .001, d = 0.72$ ). The effect was, however, significantly smaller for the blend faces than for prototype faces ( $t(43) = 5.58, p < .001, d = 0.47$ ). This pattern of results was very similar for both sets of stimuli. For set 1, participants chose samples that were 3.06 levels darker for the prototype Black faces than prototype White faces ( $t(43) = 7.17, p < .001, d = 1.27$ ), and 1.93 levels darker for blend Black faces than for blend White faces ( $t(43) = 6.01, p < .001, d = 0.87$ ). The effect for blend faces was significantly smaller than the effect for prototype faces ( $t(43) = 4.44, p < .001, d = 0.43$ ). For set 2, participants chose samples that were 2.33 levels darker for the prototype black faces than prototype white faces ( $t(43) = 5.04, p < .001, d = 0.98$ ), and 0.92 levels darker for blended Black faces than for blended White faces ( $t(43) = 3.00, p < .001, d = 0.40$ ). The effect for blended faces was significantly smaller than the effect for prototype faces ( $t(43) = 4.74, p < .001, d = 0.49$ ).

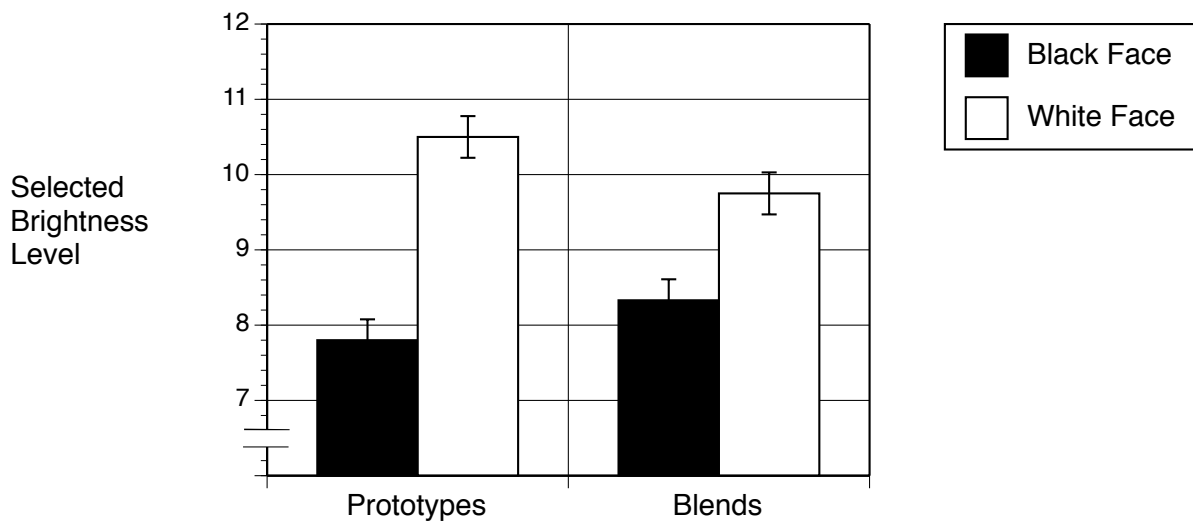


Figure 2. Brightness distortion effect in prototype and blended faces. The selected brightness level is the mean response on a 17-level scale.

We analyzed the feelings thermometer data by deriving two measures of attitudes. For the first, we subtracted ratings for Black people from ratings for White people, and for the second, we subtracted ratings for Brown (*Pardas/morenas*) people from ratings of White people, to obtain scores reflecting attitudes toward each pair of groups. Positive scores indicate relative favoritism towards White people. We then ran two regressions testing whether these scores predicted the brightness illusion (averaged across both stimulus sets). One regression included the Black vs. White attitude measure as a predictor, and one including the Black vs Brown measure as a predictor. Each regression also included participant age, sex, race (White vs. a combined category for Black/Brown), and education. For the regression testing the Blacks vs Whites feelings thermometer predictor, the feelings thermometer score was the only significant predictor ( $Beta = -.35$ ,  $t(42) = -2.05$ ,  $p = .047$ ; all other predictors  $p's > .25$ ) of the lightness effect. For the regression including the Whites relative to Browns (*Pardas/morenas*) predictor, none of the predictors were significant (all  $p's > .17$ ). Consistent with the lack of an effect of participant race on the brightness distortion effect, significant distortion effects occurred both for White participants ( $t(22) = 4.68$ ,  $p < .001$ ,  $d = 1.01$ ) and Black/Brown participants ( $t(20) = 3.83$ ,  $p = .001$ ,  $d = 0.84$ ).

## Discussion

We clearly replicated Levin & Banaji's brightness distortion effect for both sets of faces, and both in blended faces and prototype faces. Moreover, the effect occurred equally in our white and brown/black participant groups. This suggests that the distortion effect from visual race cues on perception of lightness of skin tone generalizes across cultures as well as across groups within our Brazilian sample. The effect size was comparable to that found in the original study ( $d = 1.27$  for the prototype Set 1 faces in our study – using similar stimuli, Levin & Banaji (2006) obtained effect sizes ranging between  $d = 0.75$  and  $d = 1.65$ ). This suggests that the more heterogeneous racial categories in Brazil are not accompanied by substantially different associations between face features and perceived skin color compared to the US. Our results also demonstrated that the brightness effect occurs for faces manipulated to be less different between the races, although the effect is significantly smaller in this case. In contrast to Levin & Banaji, we observed a correlation between racial attitudes and the size of the distortion effect, with a tendency for participants with less negative attitudes toward non-Whites to show higher brightness distortion. The effect was, however, not large, and went in an unexpected direction. More data were collected in Experiment 2 to further test for an attitude-brightness relationship.

### Experiment 2: Testing the effect of speech accent on brightness perception

As described above, Experiment 2 tested the novel hypothesis that, in addition to the lightness distortion caused by visual cues to race that we replicated in experiment 1, associations carried by verbal accent could have a similar effect on lightness perception. We also added a recall task that tested participants' memory for how faces were paired with races. This served partly as a manipulation check, and partly as a probe into how participants'

patterns of recall errors might reveal how they encoded the relationships between voice and race (cf. Taylor, Fiske, Etcoff, & Ruderman, 1978).

### Method

**Participants.** A total of 82 participants were recruited on campus at Universidade Federal do Rio Grande do Norte in Natal, Brazil. One participant withdrew from the experiment halfway through and was excluded from analysis, leaving 81 participants (56 female) in the final dataset for analysis. 44 participants indicated that they were white (*branca*), 32 brown (*parda*), 2 black (*preta*), 1 Asian (*amarela*), 1 indigenous (*indigena*), and one did not indicate a color. Mean age was 23.0 years (range 18-39 years).

**Apparatus.** The same computers and settings were used for presentation of the experimental stimuli as in Experiment 1. Participants wore earphones in this experiment.

**Stimuli.** The experiment used audio recordings of local speakers gathered and prepared specifically for the study. All recordings used male speakers, all of whom read aloud the same series of distinct sentences. Sentences contained neutral content, following the stimuli content of Kinzler, Dupoux, & Spelke (2007), e.g. “sometimes children play football in the park”. Recordings from four 'high-class' speakers and four 'low-class' speakers were taken and prepared as stimuli. These were collected, respectively, from highly educated university students and staff at Federal University of Rio Grande do Norte, and from working class adults at a part-time night school who were completing their fundamental education in later life. To test whether the stimulus recordings were perceived to fall into the expected social categories by our test population, we presented the voices to participants from Experiment 1 ( $n = 44$ ). After listening to the same sentence pronounced by each voice, the participants made a forced choice on two association measures: color (choice between Black and White face) and income level (choice between ‘\$’ and ‘\$\$\$\$\$’). Results are shown in Table 1.



Table 1

*Color and Income Level Associated With the Accents*

Accent	Color <sup>a</sup> (% Black)	Income <sup>b</sup> (% Poor)
<b>High-SES</b>		
Speaker A	50	34
Speaker B	41	34
<b>Low-SES</b>		
Speaker C	73	95
Speaker D	71	95

*Note.* <sup>a</sup>Numbers show the percentage of participants ( $n = 44$ ) that chose the black face when asked which face would match the speaker. <sup>b</sup>The percentage of participants that chose the '\$' symbol as association with the speaker.

As shows, for the voices of Low-SES speakers, the link with low income was virtually at ceiling, and there was also a clear association with being Black. Associations were more ambiguous for High-SES speakers. This distribution is to be expected, as more people are (very) poor than (very) rich generally so the likelihood of the high-class speakers being at ceiling levels of income is low relative to that of low-class speakers being at floor levels of income. Since the accent associations were clearly different (comparing associated color between High-SES and Low-SES,  $\chi^2(1) = 24.25, p < .001$ ; associations did not differ within speakers of High-SES or Low-SES accents,  $\chi^2(1) < 1.46, p > .23$ ) and accurately tracked the income discrepancies predicted, we proceeded to use these recordings in the experiment.

The four prototype faces from Experiment 1 were used in their racially unambiguous version, i.e. we used two prototype Black faces and two prototype White faces. Similar to Experiment 1, each face was included in five brightness variants, two progressively lighter

than the original and two progressively darker. For verbal accent, we used sound recordings of ten different sentences in Brazilian Portuguese (see above), by four different speakers. Two of these speakers were categorized as ‘high-SES’ and two as ‘low-SES’. As in Experiment 1, the experiment was programmed in PsychoPy.

**Procedure.** Participants were recruited for a study on ‘visual perception of faces’. Participants were tested in small groups, with two subjects doing the experiment simultaneously at adjacent computers. First, a research assistant entered age and sex of the participant into the computer. Participants then read instructions on the screen and began the experimental task when they were ready. In each trial, a face appeared in the middle of the screen and 0.35 seconds later a voice (ostensibly the voice of the person on the screen) spoke a sentence in Brazilian Portuguese. After the sentence ended, a grey square appeared to the right of the face. Participants adjusted the lightness of this square using the up- and down-arrows on the keyboard until they judged it to match the skin color of the face, and then pressed the spacebar to move on to the next trial. The experiment included 44 trials: 4 unrecorded practice trials followed by 40 experimental trials (10 with each prototype face). Each participant was exposed to one Black face and one White face paired with High-SES accents, and one Black face and one White face paired with Low-SES accents. The specific pairings (e.g. which Black face went with which High-SES voice) were counterbalanced between participants, and the individual trials were presented in randomized order. After the experimental tasks, participants listened again to a sentence from the experiment in each of the four voices on headphones. After listening to each voice, the participant indicated on a paper questionnaire which of the four faces had previously been paired with the voice. All participants heard the same sentence in this recall task, but the order in which they were presented was fully counterbalanced between participants. Finally, participants filled in the

feelings thermometer, additional demographic information, and the questionnaire on accent associations from Experiment 1.

## Results

We analyzed the data using a two-factor repeated-measures ANOVA with the factors race (*Black* and *White*) and accent (*High-SES* and *Low-SES*). The analysis showed a strong main effect of race,  $F(1,80) = 78.63, p < .001, r = .70$ . Participants on average chose samples 2.2 levels darker for Black faces ( $M = 8.27, SE = 0.28$ ) than for White faces ( $M = 10.46, SE = 0.26$ ), replicating the effect of race from Experiment 1. Of the 81 participants, 73 (90%) chose darker samples for Black faces than for White faces. Importantly, there was also a main effect of accent,  $F(1,80) = 6.20, p = .015, r = .27$ . Participants on average chose samples 0.43 levels darker for faces paired with a Low-SES accent ( $M = 9.15, SE = 0.26$ ) than for faces paired with a High-SES accent ( $M = 9.58, SE = 0.25$ ). There was no interaction between race and accent,  $p = .51$ .

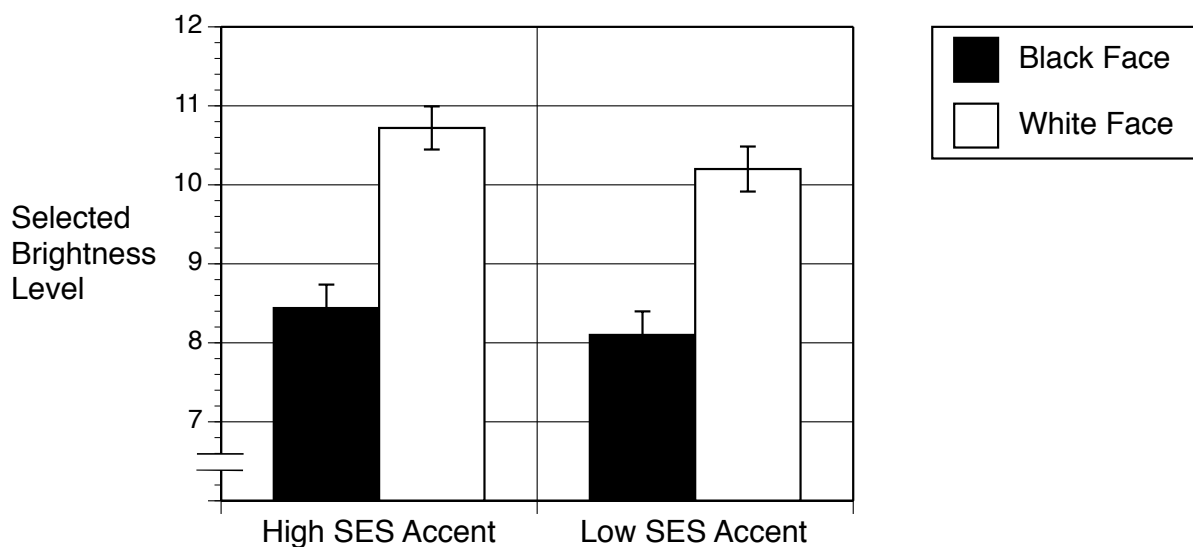


Figure 3. Impact of race and accent on brightness of chosen sample. The selected brightness level is the mean response on a 17-level scale.

As in Experiment 1 we assessed the degree to which attitudes, along with age, sex, race, education would predict the brightness distortion by running two multiple regressions

predicting the overall brightness distortion effect. One of the regressions used the difference between attitudes toward Whites and toward Blacks as the attitude measure, and one used the difference between Whites and Browns (*Pardas/morenas*). None of the predictors approached significance in either regression ( $p$ 's > .39). Consistent with the lack of an effect of participant race on the distortion effect, the distortion effect was significant both for Black/Brown participants ( $t(36) = 6.80, p < .001, d = 1.14$ ) and for White participants ( $t(43) = 5.88, p < .001, d = 0.89$ ).

**Error patterns in face-voice pairings.** Finally, participants were tested for their recall of which voice had been paired with which face. This was a difficult task for most participants, who on average got 1.5 ( $SD = 1.18$ ) out of the four pairings right (note that any one participant was able to get 0, 1, 2, or 4 pairings correct – if a participant got three pairings correct, then the remaining pairing must also be correct). To test whether participants' ability to correctly remember the voice-face pairings affected the size of the brightness distortion, we re-ran the repeated-measures ANOVA from the main analysis. This time, we added the number of pairings correctly recalled as between-groups factor (levels: 0-1 vs. 2-4 correct responses). There was no interaction between recall accuracy and the accent effect,  $p = .996$ . The recall test also allowed us to investigate if participants' error patterns might reveal how they encoded face-voice relations, similarly to the memory confusion paradigm by Taylor, Fiske, Etcoff, & Ruderman (1978) which uses errors in recall to reveal whether participants are categorizing target individuals into groups (cf. Kurzban, Tooby, & Cosmides, 2001). Within each racial category of face (Black and White), participants could make two types of mistakes: they could misattribute either an incorrect High-SES or an incorrect Low-SES speaker to any given face. If the pattern of errors were random, faces of both races would be equally likely to be misattributed to a High-SES as a Low-SES accent,

given the counterbalancing of face-voice pairings. The actual proportions of the two types of mistakes, summed across all participants, are shown below.

Table 2

*Type of Incorrect Accent Chosen as a Percentage of All Errors Made For Each Race*

Accent	White ( $n = 105$ ) <sup>a</sup>	Black ( $n = 98$ ) <sup>a</sup>
High-SES	63.8	38.1
Low-SES	36.2	61.9

*Note.* Individual participants were able to get 0, 1, 2 or 4 pairings correct. Percentages shown reflect the distribution of responses in those cases where participants got the pairing wrong.

<sup>a</sup> $N$  is the total number of times an incorrect voice was attributed to a face of this race.

Chi-square tests revealed that the error patterns were significantly different from a 50-50 distribution for both White,  $\chi^2(1) = 8.0$ ,  $p = .005$ , and Black faces,  $\chi^2(1) = 5.5$ ,  $p = .02$ . For White faces, participants were 1.8 times more likely to pick a High-SES than a Low-SES accent when they misremembered the voice; for Black faces, participants were 1.6 times more likely to pick a Low-SES than a High-SES accent when they misremembered the voice.

## Discussion

Similarly to Experiment 1, we found a strong main effect of race, again replicating Levin & Banaji's brightness distortion effect. We also found a main effect of accent, with participants choosing darker samples when faces were paired with a Low-SES accent than when they were paired with a High-SES accent. The direction of this effect matched the associations with the accents found in participants from Experiment 1, where Low-SES accents were much more likely than High-SES accents to be associated with being Black. This suggests that perception of lightness of skin tone can be influenced by information carried by speech accent, just like it can be influenced by visual cues to race. In contrast to Experiment 1, but in line with Levin & Banaji 2006, Experiment 2 did not reveal any effect

of racial attitudes on the brightness distortion effect. This suggests that the effect found in Experiment 1 was not reliable. Finally, when we analyzed participants' recall of face-voice pairings, the pattern of errors made was not random: when participants misremembered the face-voice pairing, they were more likely to err by choosing a High-SES than a Low-SES accent for White faces whereas for Black faces they were more likely to err by choosing a Low-SES than a High-SES accent. We cannot rule out an influence of demand characteristics, though the nature of the study was not revealed to participants prior to completion. Nevertheless, the pattern of errors suggests that participants' associations between race and accent biased their recall of faces and voices in the same direction as the effect found of accent on brightness perception (Low-SES accent = Black; High-SES accent = White).

## **General Discussion**

The first objective of the present work was to investigate whether the brightness distortion effect found by Levin & Banaji (2006), where people perceived the skin tone of faces to be darker when those faces had Black features compared to White features, would generalize to a Brazilian population. Experiment 1 closely followed the methods used by Levin & Banaji and replicated their result in two separate sets of faces: our participants matched darker samples to faces that had Black features than White features, with an effect size similar to Levin & Banaji (2006). Moreover, the effect was independently replicated in the different design of Experiment 2. We did not observe reliable effects of demographic variables or racial attitudes on the distortion effect. This suggests that the effect generalizes across cultures, and that the putatively more 'fluid' racial categories in Brazil do not translate into different lightness associations with the racial features used in our stimuli.

The second objective was to investigate if associations carried by speech would have a similar effect. This prediction was confirmed in Experiment 2, where we found that

participants judged faces to be darker when they were paired with a Low-SES accent than when they were paired with a High-SES accent. The effect is in line with the socio-cultural context in which Whites have higher average incomes and literacy rates than Blacks. It also accords with the data we gathered on associations with the voices used as stimuli, where voices we classified as Low-SES were dominantly associated with being Black and of low income, in contrast to the High-SES accents which had more ambiguous connotations. Our results complement those of recent studies that have contrasted accent and race as cues to social categorization (e.g. Pietraszewski & Schwartz, 2014a), and suggest that accent cues - through association - might influence categorization by race. Our findings may also inform neuroscientific approaches to face-voice integration in person perception (see e.g. Campanella & Belin, 2007 for a review), by demonstrating that statistical information about skin tone is extracted from voices and that this in turn can influence visual judgment of skin tone (Latinus & Belin, 2011).

In relation to brightness perception more specifically, our findings suggest that a broad range of categorical information may influence brightness perception, in so far that those categories are associated with expected lightness. It is important to note that these categorical influences can occur in a variety of ways. Some categorical influences might be the result of long-term learning that generates connections within perceptual (and plausibly even modular) systems (for a recent example, see Bannert & Bartels, 2013). On this view, frequently observed co-occurrence of accents and skin colors might generate an accent-brightness effect that is relatively independent of on-line activations of more abstract conceptual information. In contrast, the observed impact of accent on brightness could represent a more classic top-down effect in which accents activate abstract conceptual knowledge that in turn impacts the output of perceptual systems (Macpherson, 2012). On this

view, the perception of brightness might be more malleable and responsive to more transitory influences. This is a complex question, and it is possible both to imagine subtle variations on the above hypotheses (Vetter & Newen, 2014).

However, given that the effects of attitudes are small or non-existent in the two experiments reported here and in the original Levin and Banaji (2006) paper, we suspect that there are limits to the ability of non-perceptual processes to impact face brightness perception and that it is more likely that some form of strong perceptual instantiation (for example, actually hearing a voice or seeing a face-form that is perceptually linked to the perceived surface of the face) is necessary to produce a face-brightness effect. The interesting possibility raised by the present data is that this perceptual instantiation could be intermodal, and, as such, the findings add to the small amount of recent work documenting the impact of sound on brightness. In particular, Sutherland, Thut, & Romei (2014) demonstrate that participants perceive a grey disc to be brighter if it is associated with a looming sound, and so our finding expands the possible auditory inputs to brightness from spatial sounds such as looming to accent which may reflect an object category.

Although these are important questions, a more functional approach to brightness perception makes it clear that whatever the specific processes underlying the face brightness distortion, the effects we have observed are important because a) they generalize across populations and stimuli, b) they impact judgments that could have important social, cultural, and political consequences, and c) they reflect a form of pervasive learning in which culturally important accent information shapes at least some form of cognitive or perceptual experience. Thus, we have demonstrated that the brightness distortion is present across cultural settings, and that it pervasively influences judgments of the vast majority of individuals within those cultures in ways that may readily impact broader practices, ranging



from how different races might be visually depicted, to how they might be referred to and ultimately how they might be thought about and remembered.

## **Appendix A: Instructions for the experimental tasks**

### **Experiment 1**

#### *Instructions shown on-screen (Brazilian Portuguese)*

Você será apresentado a uma série de rostos. Cada rosto vai aparecer no centro da tela. A direita de cada rosto haverá um quadrado cinza. A sua tarefa será ajustar a cor do quadrado cinza usando as setas para cima ou para baixo até que fique semelhante ao tom de pele do rosto. Quando você terminar a tarefa, aperte a barra de espaço.

Tente fazer com que fique o mais parecido possível com o tom geral da pele do rosto, e não se preocupe se você não conseguir deixá-lo exatamente igual - basta responder com base no seu melhor julgamento.

Clique na barra de espaço quando você estiver pronto para começar.

#### *English translation of instructions*

You will be presented with a series of faces. Each face will appear in the center of the screen. A grey square will appear to the right of the face. For each face, your task is to adjust the colour of the grey square using the up- and down-arrows until the square is similar to the skin tone of the face. Indicate you have found a match by hitting the space bar.

Try your best to match the overall skin tone of the face, and don't worry if you cannot match it exactly - just respond based on your best judgment.

Hit the space bar when you are ready to begin.

### **Experiment 2**

#### *Instructions shown on-screen (Brazilian Portuguese)*

Você será apresentado a uma série de rostos. Cada rosto vai aparecer no centro da tela, e você vai ouvi-lo falar. Em seguida, um quadrado cinza aparecerá à direita do rosto. Quando o

quadrado cinza aparecer, a sua tarefa será ajustar a cor do quadrado cinza usando as setas para cima ou para baixo até que fique semelhante ao tom de pele do rosto. Quando você terminar a tarefa, aperte a barra de espaço.

Tente fazer com que fique o mais parecido possível com o tom geral da pele do rosto, e não se preocupe se você não conseguir deixá-lo exatamente igual - basta responder com base no seu melhor julgamento.

Clique na barra de espaço quando você estiver pronto para começar.

*English translation of on-screen instructions*

You will be presented with a series of faces. Each face will appear in the center of the screen, and you will hear it speak. Then a grey square will appear to the right of the face. When the grey square appears, your task is to adjust the colour of the grey square using the up- and down-arrows until the square is similar to the skin tone of the face. Indicate you have found a match by hitting the space bar.

Try your best to match the overall skin tone of the face, and don't worry if you cannot match it exactly - just respond based on your best judgment.

Hit the space bar when you are ready to begin.

### **References**

- Adelson, E. H. (2000). Lightness Perception and Lightness Illusions. In M. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (2nd ed., Vol. 3, pp. 339–351). Cambridge, MA: MIT Press.
- Bannert, M. M., & Bartels, A. (2013). Decoding the Yellow of a Gray Banana. *Current Biology*, 23(22), 2268–2272. <http://doi.org/10.1016/j.cub.2013.09.016>
- Bayard, D., Weatherall, A., Gallois, C., & Pittam, J. (2001). Pax Americana? Accent Attitudinal Evaluations in New Zealand, Australia and America. *Journal of*

*Sociolinguistics*. <http://doi.org/10.1111/1467-9481.00136>

Byrne, B., Harris, M., Consorte, J. G., & Lang, J. (1995). What's in a name? The consequences of violating Brazilian emic color-race categories in estimates of social well-being. *Journal of Anthropological Research*, 51, 389–397.

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*. <http://doi.org/10.1016/j.tics.2007.10.001>

Cohen, E. (2012). The Evolution of Tag-Based Cooperation in Humans. *Current Anthropology*, 53(5), 588–616. <http://doi.org/10.1086/667654>

Cohen, E., & Haun, D. (2013). The development of tag-based cooperation via a socially acquired trait. *Evolution and Human Behavior*, 34(3), 230–235.  
<http://doi.org/10.1016/j.evolhumbehav.2013.02.001>

Cornsweet, T. N. (1970). *Visual Perception*. New York: Academic Press.

Degler, C. N. (1971). *Neither Black Nor White. Slavery and Race Relations in Brazil and the United States*. Madison: University of Wisconsin Press.

Dehghani, M., Khooshabeh, P., Nazarian, A., & Gratch, J. (2014). The Subtlety of Sound: Accent as a Marker for Culture. *Journal of Language and Social Psychology*, 1–20.  
<http://doi.org/10.1177/0261927X14551095>

Fish, J. M. (2005). Mixed Blood. In J. P. Spradley & D. W. McCurdy (Eds.), *Conformity and conflict: Readings in cultural anthropology* (pp. 249–259). New York: Allyn & Bacon.

Foon, A. E. (2001). A Social Structural Approach to Speech Evaluation. *The Journal of Social Psychology*, 126(4), 521–530.

Fuertes, J. N., Potere, J. C., & Ramirez, K. Y. (2002). Effects of speech accents on interpersonal evaluations: implications for counseling practice and research. *Cultural*

- Diversity & Ethnic Minority Psychology*, 8, 346–356. <http://doi.org/10.1037//1099-9809.8.4.346>
- Gil-White, F. J. (2001). Sorting is not Categorization: A Critique of the Claim that Brazilians Have Fuzzy Racial Categories. *Journal of Cognition and Culture*, 1(3), 219–249.
- Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., ... Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, 106(4), 795–834. <http://doi.org/10.1037//0033-295X.106.4.795>
- Giles, H. (Ed.). (2012). *Handbook of Intergroup Communication*. New York: Routledge.
- Harris, M., Consorte, J. G., Lang, J., & Byrne, B. (1993). Who Are the Whites - Imposed Census Categories and the Racial Demography of Brazil. *Social Forces*, 72(2), 451–462. <http://doi.org/10.2307/2579856>
- Harris, M., Consorte, J. G., Lang, J., & Byrne, B. (1995). A reply to Telles. *Social Forces*, 73(4), 1609–1611.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223–233.
- Helmholtz, H. von. (1866). *Helmholtz' treatise on physiological optics*. New York: Optical Society of America.
- Jones, D. (2009). Looks and Living Kinds: Varieties of Racial Cognition in Bahia, Brazil. *Journal of Cognition and Culture*, 9(3), 247–269.
- Kayser, A. (1997). *Heads*. Abbeville Press.
- Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, 104(30), 12577–12580. <http://doi.org/10.1073/pnas.0705345104>

- Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent Trumps Race in Guiding Children's Social Preferences. *Social Cognition*.  
<http://doi.org/10.1521/soco.2009.27.4.623>
- Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences of the United States of America*, 98(26), 15387–15392. <http://doi.org/10.1073/pnas.251541498>
- Latinus, M., & Belin, P. (2011). Human Voice Perception. *Current Biology*, 21(4), R143–R145.
- Levin, D. T., & Banaji, M. R. (2006). Distortions in the perceived lightness of faces: the role of race categories. *Journal of Experimental Psychology: General*, 135(4), 501–12.  
<http://doi.org/10.1037/0096-3445.135.4.501>
- Lotto, R. B., & Purves, D. (2002). The empirical basis of color perception. *Consciousness and Cognition*, 11(4), 609–629. [http://doi.org/10.1016/S1053-8100\(02\)00014-4](http://doi.org/10.1016/S1053-8100(02)00014-4)
- MacLin, O. H., & Malpass, R. S. (2003). The ambiguous-race face illusion. *Perception*, 32(2), 249–252.
- MacLin, O., & Malpass, R. (2001). Racial Categorization of Faces. The Ambiguous Race Effect. *Psychology, Public Policy, and Law*, 7(1), 98–118.
- Macpherson, F. (2012). Cognitive Penetration of Colour Experience: Rethinking the Issue in Light of an Indirect Mechanism. *Philosophy and Phenomenological Research*, 81(1), 24–62.
- Mitterer, H., & de Ruiter, J. P. (2008). Recalibrating color categories using world knowledge. *Psychological Science*, 19(7), 629–634.
- Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience*

*Methods*, 162(1-2), 8–13.

Peterson, M. (1994). Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science*, 3(4), 105–111.

Pietraszewski, D., & Schwartz, A. (2014a). Evidence that accent is a dedicated dimension of social categorization, not a byproduct of coalitional categorization. *Evolution and Human Behavior*, 35, 51–57. <http://doi.org/10.1016/j.evolhumbehav.2013.09.005>

Pietraszewski, D., & Schwartz, A. (2014b). Evidence that accent is a dimension of social categorization, not a byproduct of perceptual salience, familiarity, or ease-of-processing. *Evolution and Human Behavior*, 35, 43–50.  
<http://doi.org/10.1016/j.evolhumbehav.2013.09.006>

Purves, D., Williams, S. M., Nundy, S., & Lotto, R. B. (2004). Perceiving the intensity of light. *Psychological Review*, 111, 142–158. <http://doi.org/10.1037/0033-295X.111.1.142>

Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, 100, 16–29. <http://doi.org/10.1037/a0021522>

Schwartzman, L. F. (2007). Does Money Whiten? Intergenerational Changes in Racial Classification in Brazil. *American Sociological Review*, 72, 940–963.

Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science*, 4(6), 379–384. <http://doi.org/10.1111/j.1467-9280.1993.tb00585.x>

Sutherland, C. A. M., Thut, G., & Romei, V. (2014). Hearing brighter: Changing in-depth visual perception through looming sounds. *Cognition*, 132(3), 312–323.  
<http://doi.org/10.1016/j.cognition.2014.04.011>

- Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*. <http://doi.org/10.1037/0022-3514.36.7.778>
- Telles, E. E. (2004). *Race in Another America: The Significance of Skin Color in Brazil*. Princeton, NJ: Princeton University Press.
- Telles, E. E., & Sue, C. A. (2009). Race Mixture: Boundary Crossing in Comparative Perspective. *Annual Review of Sociology*, 35, 129–146.
- Vetter, P., & Newen, A. (2014). Varieties of cognitive penetration in visual perception. *Consciousness and Cognition*, 27(October 2015), 62–75.  
<http://doi.org/10.1016/j.concog.2014.04.007>