



OPEN

## Evidence against the “anomalous-is-bad” stereotype in Hadza hunter gatherers

Clifford I. Workman<sup>1,2,3,6</sup>, Kristopher M. Smith<sup>4,6</sup>, Coren L. Apicella<sup>5</sup> & Anjan Chatterjee<sup>1,2,3</sup>

People have an “anomalous-is-bad” stereotype whereby they make negative inferences about the moral character of people with craniofacial anomalies like scars. This stereotype is hypothesized to be a byproduct of adaptations for avoiding pathogens. However, evidence for the anomalous-is-bad stereotype comes from studies of European and North American populations; the byproduct hypothesis would predict universality of the stereotype. We presented 123 Hadza across ten camps pairs of morphed Hadza faces—each with one face altered to include a scar—and asked who they expected to be more moral and a better forager. Hadza with minimal exposure to other cultures chose at chance for both questions. Hadza with greater exposure to other cultures, however, expected the scarred face to be less moral and a better forager. These results suggest the anomalous-is-bad stereotype may be culturally shared or learned erroneously through associations with population-level differences, providing evidence against a universal pathogen avoidance byproduct hypothesis.

If Hollywood is to be believed, then people are usually only as *good* as they are good looking. In contrast to heroes, movie villains are more frequently depicted with visible facial anomalies, like scars, that negatively impact perceptions of attractiveness<sup>1</sup>. In the *Star Wars* universe, for instance, Anakin Skywalker’s transformation into Darth Vader is marked by the simultaneous corruption of his moral character and of his physical form. The notion that facial appearance weighs on character is more than just a familiar storytelling device—participants from primarily English-speaking countries have a “beauty-is-good” stereotype and are more likely to ascribe positive personal qualities to beautiful people<sup>2–5</sup>. On the other hand, US students and online workers have an “anomalous-is-bad” stereotype and are more likely to ascribe negative qualities to people with deviations from prototypically attractive features (e.g., craniofacial anomalies), especially negative attributions of moral character<sup>5–8</sup>.

People may use craniofacial anomalies and other deviations from facial typicality to infer moral character because these anomalies are perceived as cues to infectious disease. Facial characteristics associated with beauty on the other hand, like symmetry and averageness, may serve as cues of healthiness that prospective partners can use to assess mate quality<sup>9</sup>. US students with more attractive faces have better in vitro immunological responses to bacterial and viral infections<sup>10</sup>. Consistent with the foregoing evolutionary account, preferences for symmetry and averageness appear to be universal. The Hadza hunter-gatherers of Tanzania show preferences for symmetrical Hadza and European faces and for more average Hadza faces (though—perhaps owing to a lack of familiarity—not for more average European faces)<sup>11,12</sup>. Of note, although some prior work has suggested a more specific link between facial symmetry and pathogen avoidance, a recent registered report casts doubt on such a link<sup>13</sup>. Humans have an evolved behavioral immune system to identify and avoid potential vectors of pathogens, using cues that were historically associated with increased pathogen risk, such as coughs, sneezes, lesions, and bodily fluids<sup>14</sup>. Failing to detect threats of pathogens can have costly consequences, whereas vigilance against such threats is likely to be less costly, even when triggered erroneously by irrelevant cues<sup>15–17</sup>. In conditions of uncertainty, the behavioral immune system errs on the side of caution by biasing behavior towards vigilance in order to limit the harmful consequences of potential errors<sup>16,17</sup>.

<sup>1</sup>Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>2</sup>Penn Brain Science Center, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>3</sup>Penn Center for Neuroaesthetics, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>4</sup>Department of Anthropology, Washington State University, Pullman, WA 99163, USA. <sup>5</sup>Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA. <sup>6</sup>These authors contributed equally: Clifford I. Workman and Kristopher M. Smith. ✉email: cliffworkmanphd@gmail.com; kristopher.m.smith@wsu.edu



**Figure 1.** A sample set of composite male faces used in the experiment (face set M1, condition A). The complete set of stimuli are available at <https://osf.io/eqftk/>.

Cues of pathogen threat may include facial anomalies, even when such anomalies are acquired independently of any underlying illnesses<sup>18</sup>. Indeed, facial anomalies and unattractive faces are treated as expressions of infectious disease that elicit disgust and avoidance in perceivers<sup>19–21</sup>. Moreover, expressions of negative bias in attitudes and behavior towards people with facial anomalies are predicted by disgust-sensitive brain responses to anomalous faces<sup>6–8</sup>, further evidence that disgust underpins reactions to anomalous faces and that people erroneously infer pathogen threat from anomalies<sup>18</sup>. Abundant cross-cultural evidence suggests that feelings of disgust motivate the avoidance of both pathogen threats and moral threats<sup>22–27</sup>. People express and report disgust toward pathogen-irrelevant actions, such as unfairness, dishonesty, violence, and child abuse<sup>28–32</sup>, and stronger experiences of disgust toward violations are associated with increased moralistic punishment<sup>23,25,33</sup>. Children as young as 6, for instance, reportedly responded with disgust more often for moral violations (e.g., meanness directed at another person) than for negatively valenced yet morally neutral actions (e.g., watching a sad movie with another person)<sup>25</sup>. Finally, people also infer negative moral character from experiences of disgust, which motivates the avoidance of immoral others<sup>34,35</sup>.

If the anomalous-is-bad stereotype is a byproduct of adaptations designed to detect and avoid potential pathogen threats, then it should be expressed broadly across cultures, which requires testing in diverse samples<sup>36</sup>. However, most of the evidence for the stereotype comes from Western populations, which are outliers on many psychological dimensions compared to the global population<sup>37,38</sup>. Existing evidence on perceptions of trustworthiness from facial appearance suggests variation of the stereotype between and within cultures. For example, Dutch<sup>39</sup>, US, and Japanese<sup>40</sup> adults vary in the extent to which they believe traits can be inferred from facial features (i.e., lay physiognomy beliefs), and people with stronger beliefs in physiognomy are more likely to form character impressions using facial appearance. Further, the extent to which people infer trustworthiness and dominance from facial appearance varies across cultures<sup>41</sup>. While these studies did not examine the anomalous-is-bad stereotype cross-culturally, they provide evidence that specific impressions formed from facial appearance are not universal.

Here, we examine the anomalous-is-bad stereotype among the Hadza. The Hadza are not isolated, and have had contact since at least the nineteenth century with neighboring populations, including Datoga, Isanzu, Sukuma, and Maasai<sup>42</sup>; however, they have had minimal contact with Western media, such as television, magazines, and cinema. In addition, the Hadza practice intentional scarification, making small vertical or lateral incisions on the cheeks with a small knife. The function of these incisions is thought to be manifold: they serve the aesthetic purpose of marking children as “Hadza,” and they allow for the application of medicines when treating injuries and ailments<sup>42–44</sup>. Intentional scarification for purposes of beautification and rites of practice are common in East Africa<sup>45</sup>. Intentional scarification appears to place limits on claims about the universality of the anomalous-is-bad stereotype. Scars acquired *intentionally* may function as signals of robust health and pathogen resistance that enhance attractiveness<sup>46</sup> with no such benefit conferred to unintentional scars. As such, the Hadza may not draw the same inferences from scarred faces as Western participants. We asked participants to choose between pairs of faces—one with and one without a superimposed scar (see Fig. 1)—and asked them

Face set	Condition	Good forager	Good heart
F1	A	0.39	0.34
	B	0.59	0.54
F2	A	0.52	0.28
	B	0.50	0.59
M1	A	0.63	0.45
	B	0.54	0.43
M2	A	0.60	0.39
	B	0.54	0.52
Total		0.54	0.43

**Table 1.** Proportion of responses choosing anomalous face by face set, condition, and trait. Numeric values are the proportion of responses that chose the anomalous face for that trait. Face sets with women are notated with “F” and face set with men are notated with “M.” The only difference between conditions was which face had the photoshopped scar.

Parameter	Median	90% HDI	<i>pd</i>	% in ROPE
Good forager	0.25	− 0.03, 0.54	0.932	33.1
Good heart	− 0.27	− 0.59, 0.05	0.921	29.6
Good forager × Exposure	0.13	− 0.03, 0.30	0.910	68.3
Good heart × Exposure	− 0.13	− 0.29, 0.04	0.913	69.7

**Table 2.** Parameter estimates from the marginal posterior distributions of the model. The point estimate is the median value of the marginal posterior distribution. The 90% highest density interval (HDI) is the smallest interval that contains 90% of the posterior distribution, *pd* is the proportion of the posterior in the same direction as the median, and % in ROPE is the proportion of the posterior distribution falling between the region of practical equivalence (− 0.18).

to indicate which face appeared more moral and which appeared to be a better forager. If the Hadza are no less likely to choose the scarred face, then this would be evidence that the anomalous-is-bad stereotype is not universal. We ask about morality and foraging ability because some evidence suggests that the stereotype is specific to morality<sup>5</sup>, and foraging ability is an important dimension on which Hadza evaluate campmates<sup>47,48</sup>. Asking about foraging ability thereby allows us to examine the specificity of the stereotype. Finally, we also explore variation within the Hadza as a function of exposure to non-Hadza cultures. Exposure to non-Hadza cultures is associated with economic biases<sup>49</sup>, emotional closeness to other ethnic groups<sup>50</sup>, greater sharing in an economic game<sup>51</sup>, and preferring to live with more generous campmates<sup>47</sup>. Exposure may similarly moderate whether and how they infer character qualities about other Hadza.

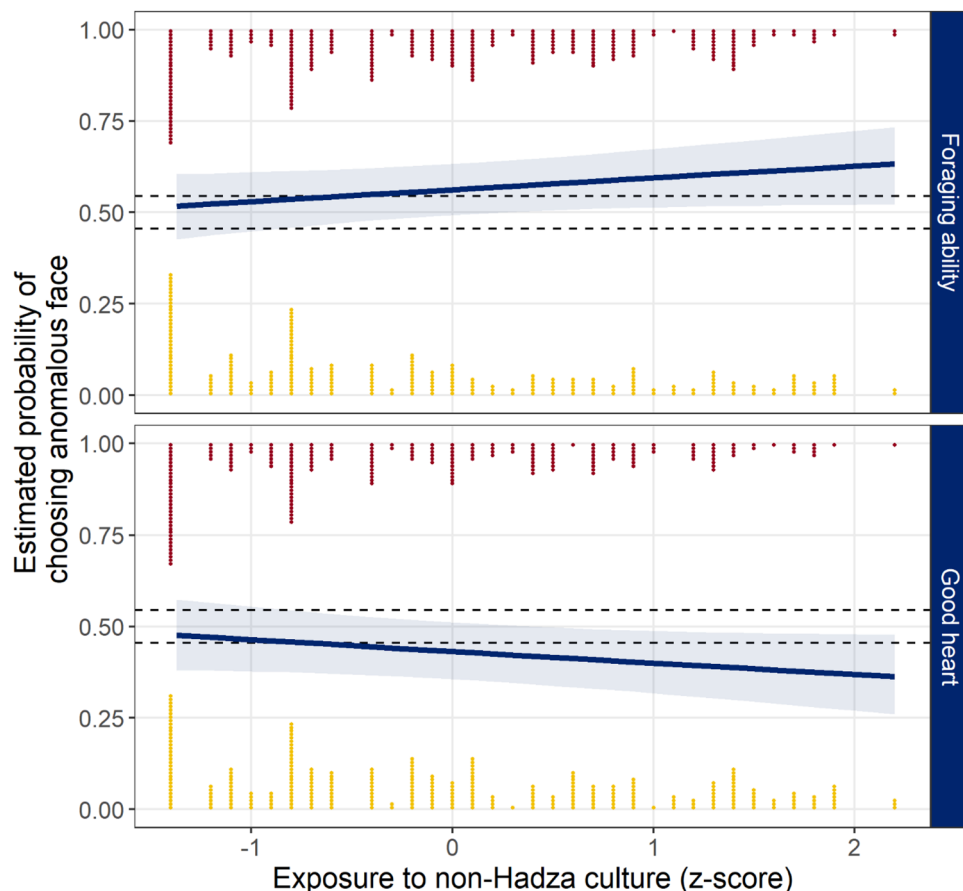
## Results

Table 1 presents the proportion of responses choosing the anomalous face by face set, condition, and trait. The only difference between conditions was which face in the pair had the facial anomaly (see Method for details). Before presenting estimates from the model, two patterns are noticeable in the raw data. First, collapsed across all face sets, participants were more likely to choose the anomalous face for good forager, and this pattern was observed within all but one face set. Second, collapsed across all face sets, participants were less likely to choose the anomalous face for good heart, and this pattern was observed in five of the eight face sets and conditions.

Table 2 presents the estimates from the marginal posterior distribution for each population parameter. When considering who is a better forager, the probability of preferring the anomalous face was positive; further, the effect of exposure on the probability of choosing the anomalous face as a better forager was positive. When considering who has a better heart, the probability of preferring the anomalous face was negative; further, the effect of exposure on the probability of choosing the anomalous face as having a better heart was negative.

We next examined the joint posterior prediction distribution and whether, across the population, participants preferred the non-anomalous face for each trait, collapsing across participant exposure, face, and condition. The probability of preferring the anomalous face as a better forager was greater than chance, *Med.* = 0.57, 90% HDI: 0.48–0.68, *pd* = 0.899. The probability of preferring the anomalous face as having a better heart was less than chance, *Med.* = 0.42, 90% HDI: 0.31–0.52, *pd* = 0.900.

Finally, we examined the joint posterior prediction distribution by participants' exposure to other cultures. Examining participants with the lowest exposure, the probability of preferring the anomalous face as a better forager was about chance, *Med.* = 0.52, 90% HDI: 0.43–0.61, *pd* = 0.630, 57.3% in ROPE, and similarly the probability of preferring the anomalous face as having a better heart was about chance, *Med.* = 0.48, 90% HDI: 0.38–0.57, *pd* = 0.660, 53.0% in ROPE. However, for participants with the highest exposure, the probability of preferring the anomalous face as a better forager was greater than chance, *Med.* = 0.63, 90% HDI: 0.53–0.74,



**Figure 2.** Joint posterior predictions by participants' exposure to non-Hadza culture as a z-score and trait being evaluated. Lines are the median estimate from the posterior and shaded regions are the 90% highest density intervals (HDI). Points are individual responses (for plotting, exposure values were rounded to the nearest tenth), with red points at the top indicating when the participant chose the anomalous face, and the yellow points at the bottom indicating when the participant chose the non-anomalous face. Dashed lines indicate the region of practical equivalence (ROPE).

$pd = 0.972$ , 8.9% in ROPE, whereas the probability of preferring the anomalous face as having a better heart was less than chance,  $Med. = 0.36$ , 90% HDI: 0.26–0.47,  $pd = 0.972$ , 8.8% in ROPE. Figure 2 presents the estimated probabilities as a function of exposure to non-Hadza culture for each trait.

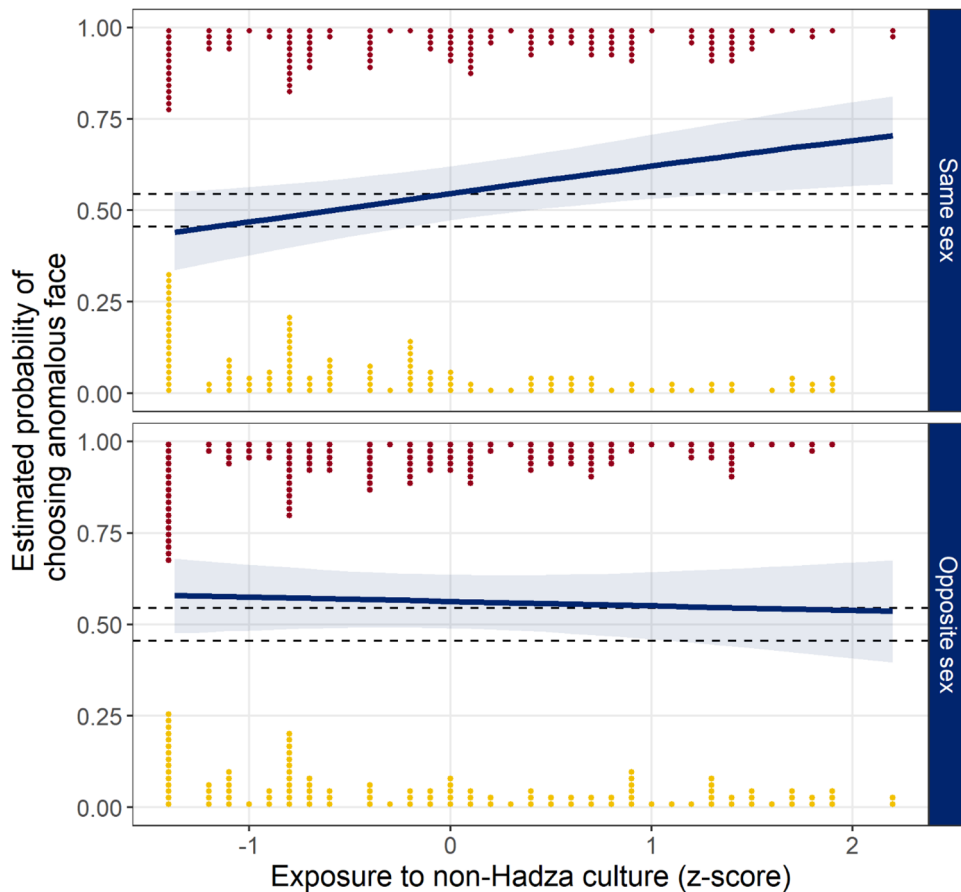
**Analyses by same-sex and opposite-sex judgments.** We analyzed the data in a second model including an interaction with whether the participant was judging a face that was the same-sex or opposite-sex. See Table 3 for parameter estimates from the model. When judging which face has a better heart, participants with greater exposure were more likely to choose the non-anomalous face, regardless of whether the judgments were of same- or opposite-sex faces. However, when judging which face is a better forager, preferences differed between same- and opposite-sex judgments. When judging faces of the same-sex, participants with greater exposure were more likely to choose the anomalous face as a better forager, but when judging faces of the opposite-sex, participants were slightly more likely to choose the anomalous face as a better forager, regardless of exposure (see Fig. 3 for predicted probabilities of choosing anomalous faces for good forager judgments).

## Discussion

People in English-speaking populations have a beauty-is-good stereotype, inferring that people perceived as more attractive are more moral and that people perceived as physically unattractive are less moral<sup>3,5</sup>. People with deviations from facial typicality, such as craniofacial anomalies like scars, are especially likely to be viewed as morally worse, whether they are perceived as attractive or not. In other words, people have an anomalous-is-bad stereotype<sup>6</sup>. We found evidence that this stereotype is not culturally universal. Hadza hunter-gatherers with minimal exposure to other cultures were not more likely to think other Hadza with facial scarring were less moral or worse foragers. However, Hadza who regularly interact with outside cultural groups were more likely to think Hadza with facial scarring were less moral and were better foragers. These results suggest the

Parameter	Median	90% HDI	<i>pd</i>	% in ROPE
<b>Same-sex judgments</b>				
Good forager	0.19	- 0.12, 0.48	0.852	46.8
Good heart	- 0.21	- 0.54, 0.14	0.846	41.2
Good forager × Exposure	0.31	0.07, 0.53	0.989	17.8
Good heart × Exposure	- 0.13	- 0.35, 0.09	0.830	63.9
<b>Opposite-sex judgments</b>				
Good forager	0.25	- 0.05, 0.55	0.921	32.4
Good heart	- 0.30	- 0.65, 0.02	0.926	26.4
Good forager × Exposure	- 0.05	- 0.27, 0.17	0.640	79.3
Good heart × Exposure	- 0.12	- 0.37, 0.08	0.823	64.4

**Table 3.** Parameter estimates from the marginal posterior distributions of the model analyzing same- and opposite-sex judgments. The point estimate is the median value of the marginal posterior distribution. The 90% highest density interval (HDI) is the smallest interval that contains 90% of the posterior distribution, *pd* is the proportion of the posterior in the same direction as the median, and % in ROPE is the proportion of the posterior distribution falling between the region of practical equivalence (-0.18).



**Figure 3.** Joint posterior predictions of choosing the anomalous face for good forager judgments by participants’ exposure to non-Hadza culture as a z-score and whether the face was same- or opposite-sex. Lines are the median estimate from the posterior and shaded regions are the 90% highest density intervals (HDI). Points are individual responses (for plotting, exposure values were rounded to the nearest tenth), with red points at the top indicating when the participant chose the anomalous face, and the yellow points at the bottom indicating when the participant chose the non-anomalous face. Dashed lines indicate the region of practical equivalence (ROPE).



anomalous-is-bad stereotype is culturally shaped, providing evidence against the hypothesis that the stereotype is a universal byproduct of the behavioral immune system.

Cultural exposure may moderate character inferences among the Hadza through the acquisition of beliefs from neighboring populations about what constitutes facial attractiveness. Previous research finds that culture influences attractiveness judgments. For example, South African Zulu migrants to the UK rated women with a lower body-mass index and waist-to-hip ratio as more attractive compared to Zulus residing in South Africa<sup>52</sup>. South Korean students with greater exposure to US media are more likely to rate thin models as attractive<sup>53</sup>. Hadza find more average Hadza faces more attractive, but not more average white European faces, suggesting exposure to different faces is necessary to judge average faces as more attractive<sup>12</sup>. Some Hadza may have learned from neighboring populations that facial scars are unattractive and consequently infer negative character traits from facial scarring. Alternatively, Hadza could believe facial scarring is unattractive but, through exposure to non-Hadza culture, acquire beliefs linking specific traits with attractiveness. US, Japanese, and Dutch participants who do not believe traits can be inferred from facial features are less likely to make such inferences<sup>39,40</sup>, providing direct evidence that explicit beliefs can influence the social evaluation of faces. Some Hadza may similarly learn associations between moral character and attractiveness from neighboring populations. One possibility is that Hadza with exposure to non-Hadza culture learn that facial scars bear negatively on perceptions of attractiveness, which then has downstream negative consequences for judgments of moral character, instead of learning that scars bear directly on moral character. This is an interesting prospect, but is at odds with the finding in this study that scars also have positive consequences for judgments of foraging ability, as better foragers are considered more attractive<sup>54</sup>.

Exposure to non-Hadza cultures could also moderate the moral character concept among Hadza. In Western populations, intent to do harm or to commit a moral violation is central to moral character judgments<sup>55,56</sup>. Intention to do wrong suggests a desire to commit the act, which is used to infer the potential to commit moral violations again in the future. That is, intention suggests the moral character concept is stable among Western participants<sup>57</sup>. However, there is variation across populations in the extent to which intentions influence moral judgments, and many small-scale populations, such as the Hadza, Himba, and Yesawa, do not condemn intentional actions more than unintentional actions<sup>58,59</sup>. Further, Hadza disagree on perceptions of who is more moral<sup>48</sup> and do not infer moral character from others' decisions in response to sacrificial dilemmas<sup>60</sup>, suggesting they may not think morality is a stable characteristic of a person. If so, then moral character would not be associated with other relatively stable traits, such as facial appearance. However, as Hadza become more exposed to other cultures, their moral character concept may change, rendering them more willing to infer character from physical characteristics.

Alternatively, culture may play a less direct role in forging associations between scars and perceptions of moral character. It is possible that, instead of internalizing the “anomalous-is-bad” stereotype, Hadza with greater exposure to non-Hadza culture may be perceived as more prestigious among the Hadza, having greater education and access to markets. Hadza may also infer that people with unintentional scars are spending more time in the bush and less time in villages. Hadza have higher levels of fluctuating asymmetry compared to European and North American populations<sup>61</sup>, possibly because of high rates of accidents Hadza experience while foraging<sup>44,62</sup>. If true, then Hadza may also infer that people with scars have less exposure and lower status, and Hadza with greater exposure—and thus greater prestige—negatively judge Hadza with anomalous faces as lower status. Across cultures, valuations of social and moral standing are correlated<sup>63</sup>, and Hadza with greater exposure may also judge lower-status Hadza as less moral. We do not have the data to explore this hypothesis in the current study, though addressing this and the alternatives presented above more directly is a goal of future research.

Surprisingly, participants with greater exposure to non-Hadza culture thought Hadza with scars were better foragers. Prior studies report that people make negative attributions from physical unattractiveness for a broad range of traits, including competence<sup>3,8</sup>, whereas others report effects specific to moral traits<sup>5</sup>. This finding could be an experimental artifact because foraging judgments were always asked after good heart judgments. Participants could have avoided selecting the same faces for both traits, such that picking the unscarred face for the good heart judgment would likely be followed by picking the scarred face for foraging. However, in previous studies, Hadza do not have difficulty discriminating between pairs of faces presented multiple times<sup>11,12</sup>. The significance of exposure to non-Hadza culture for judgments of foraging ability is further complicated by the sexes of the raters and of the faces being judged. Similar to what was observed without accounting for sex, participants with more exposure were more likely to think same-sex anomalous faces belonged to better foragers but, regardless of exposure, all participants chose the anomalous opposite-sex faces as better foragers. Consistent with evidence that facial scars may serve as indicators of beauty in some small-scale African societies, the presence of scars on opposite-sex Hadza faces may have positively influenced their perceived attractiveness with especially potent consequences for foraging judgments<sup>45,54</sup>. If so, however, such a positive effect would have also been expected for moral character judgments instead of the negative effect we observed. Another explanation is that accidents and injuries happen somewhat frequently during foraging bouts<sup>44</sup>. Hadza with facial scars may look more “bush-like” to other Hadza, who then interpret scars as indicators of foraging ability. Further research is needed to unpack the significance of exposure, sex, and attractiveness to judgments of foraging ability.

Do we assume people with facial anomalies are villains because they are portrayed as such in movies, or are villains portrayed with facial anomalies because of a pre-existing association between morality and beauty? Our data suggest that culture shapes the inferences people draw from physical appearance. Specifically, we find that the anomalous-is-bad stereotype is not expressed universally and is largely absent among Hadza without exposure to other cultures. However, there is variation within the Hadza as a function of their exposure. This study provides evidence against the hypothesis that the anomalous-is-bad stereotype is a byproduct of adaptations to avoid pathogens. The contrast between these findings and earlier work that detected the anomalous-is-bad

stereotype in Western populations highlights the importance of including more diverse populations in person perception research.

## Method

**Population.** The Hadza are a population of nomadic hunter-gatherers living in northwestern Tanzania. About 1,000 people today speak Hadzane and identify as Hadza, but only approximately 300 still obtain a majority of their calories from foraged goods<sup>44,62</sup>. Hadza live in temporary residence bands called camps, with about 20–30 adults and children in each camp. Camps usually consist of about two to three unrelated nuclear families<sup>62</sup>. Camps move location every six to eight weeks as local resources are consumed, and any individual can move freely to another camp at any time. Each year, individual Hadza interact with about one-fifth of the foraging population<sup>64</sup>, and from year to year, Hadza live with approximately one-in-five of their previous campmates<sup>65</sup>. The Hadza are highly interdependent with their campmates, and all food brought back to camp—especially meat—is shared with all other members of the camp<sup>44</sup>. Like other hunter-gatherer populations, the Hadza have a sexual division of labor, in which men hunt for large game and collect honey, while women collect plant goods such as tubers and berries<sup>44</sup>.

**Sample.** KMS and a team of research assistants visited 10 camps in September and October 2019. Camps were sampled using a snowball sampling procedure—after finishing interviews at one camp, someone in the camp would direct the research team to the next nearest camp. There was no predetermined sample size. The team sampled camps until no more camps could be located, and interviewed 126 adults, with three exclusions for cognitive impairments that prevented participants from understanding the task. The final sample was 123 adults (60 women, 89 married) with a mean age of  $M = 37.6$  ( $SD = 14.2$ ) years. Each participant made eight decisions in the experimental task for a total of 984 observations in our analyses.

**Procedure.** *Forced-choice task.* Participants were presented with four pairs of same-sex faces that were composite photographs representing novel Hadza. In each pair, one face had visible scars while the other had no scarring. After looking at each pair, participants chose the person they believed better exemplified a given trait. Participants were first asked which person in the photograph had a better heart, which is a judgment of global moral character<sup>48</sup>. Participants indicated their choice by pointing at the photograph. After answering for all four pairs, participants were again presented each pair and asked which man was a better hunter, and which woman was a better gatherer. Participants always saw the pairs in the same order, with the two male pairs presented first. The questions were also asked in the same order each time. Before their interviews, participants were randomly assigned to a condition by coin flip. The only difference between conditions was which face in each pair had visible scarring. This part of the interview was conducted in Hadzane by a Hadza research assistant and overseen by a Tanzanian research assistant.

*Survey of exposure to other cultures.* We estimated the extent to which each participant had been exposed to non-Hadza cultures using their knowledge of and experience with the surrounding cultural groups. Specifically, participants reported the number of years they attended school, whether they had worked a job, how high they could count in Swahili, if they had visited Arusha or lived outside of Hadzaland, if they knew the capital of Tanzania, and whether they knew the Tanzanian president, Barack Obama, Nelson Mandela, and Mahatma Gandhi. A two-factor structure best fit the data with responses to all but the questions about international figures loading better on to the first factor. A standardized composite score was computed using the weighted sums from the first factor as a measure of exposure to other cultures<sup>47</sup>. This part of the interview was conducted in Swahili by a Tanzanian research assistant and overseen KMS.

**Stimuli.** The experimental stimuli were constructed using 2D headshot photographs acquired from actual Hadza as part of an earlier study<sup>48</sup>. Consenting adult Hadza were photographed in 2016 with a Fujifilm Instax Mini 90 Classic Instant Film Camera. To standardize the photographs, participants were—whenever possible—positioned in front of a portable screen and instructed to adopt an emotionally neutral facial expression.

A total of 32 face photographs (16 male, 16 female) were pre-processed using established procedures<sup>6,66</sup>. First, the photographs were normalized to inter-pupillary distance with algorithms from the OpenCV computer vision library (<https://opencv.org/>) and facial landmarks from the dlib machine learning toolkit (<http://dlib.net/>). Second, they were resized and cropped in IrfanView (<https://irfanview.com/>; width: 380px; height: 570px) and then edited automatically (remove.bg, <https://www.remove.bg/>) and then manually (GIMP 2, <https://www.gimp.org/>) to remove backgrounds.

After pre-processing, each of the 32 face photographs was submitted to the InterFace software package, used to place landmarks at 82 fiducial points on each face<sup>67</sup>. The resulting coordinates were converted to JPsychoMorph format<sup>68</sup>. Then, JPsychoMorph's averaging feature was used to generate 16 composite images (8 male, 8 female) derived from subsets of 4 male or female faces from the complete set of 32 photographs. This required a total of 64 photographs (16 composites  $\times$  4 faces), so all 32 photographs were used twice (always in different composites). Next, visual inspection was used to identify and eliminate 8 composites that looked unrealistic owing to averaging errors. This resulted in a set of 8 useable composites (4 male, 4 female).

Finally, copies of the 8 composites were manually edited in Adobe Photoshop. Specifically, images of real bodily scars identified through web searches were carefully superimposed onto copies of all 8 composites. The final set of stimuli was 16 composites of four male and female faces, with an anomalous and non-anomalous version of each.

**Data analysis and computation.** We analyzed participants' decisions in Bayesian multilevel regression models<sup>69</sup>. We modeled the choice as a Bernoulli function with probability  $p$  of choosing the anomalous face;  $\text{logit}(p)$  was a function of the trait the participant was being asked about, the participant's exposure to other cultures and its interaction with the trait in question. We specified the model not to estimate an intercept and instead estimated each effect as a difference from responding at chance, with separate estimates for each trait. We included varying slopes of each trait for participant and camp. In addition, for face set, we included varying slopes for each condition, each trait, and their interactions. This approach effectively estimates which face for each set was more likely to be picked for each trait. All varying effects were modeled as multivariate normal distributions. Priors were weakly regularizing priors centered around zero. Prior predictive checks verified that all but the most extreme of values (e.g., choosing the anomalous or non-anomalous face with >95% probability) were plausible under our priors.

In the analyses, we present marginal posterior distributions of the population parameters and posterior distributions of predicted probabilities. For the distribution of predicted probabilities, we compute expected probability of choosing the anomalous face from trait, exposure, face set, and condition. We then collapse estimates across variables not of interest for the analysis to get a marginal posterior of the predicted values. We use the median of the distribution as a measure of the central tendency, 90% highest density intervals (HDI) as a measure of uncertainty, probability of direction ( $pd$ ) as a measure of existence of an effect, and percent of the posterior in the region of practical equivalence (ROPE) as a measure of significance<sup>70</sup>. HDI is the smallest continuous interval of values that 90% of the posterior falls between,  $pd$  is the proportion of the posterior above (or below depending on the sign of the median) zero (for parameter values) or 0.50 (for predicted probabilities), and % in ROPE is the proportion of the posterior distribution that falls within a predefined region around zero (or 0.50). Existence is the probability there is an effect in the direction indicated by the median, whereas significance (as indicated by the inverse of the percent in ROPE) is the probability that the effect is practically different from zero<sup>71</sup>. We defined the ROPE as the interval of  $-0.18$ – $0.18$  following guidelines for a “small” effect for standardized coefficients in logistic regression<sup>70,72</sup>.

We conducted the analyses in R (4.1.1), using the brms package (2.15.0)<sup>73</sup> with cmdstanr (0.4.0)<sup>74</sup> to fit the model, the tidyverse (1.3.1)<sup>75</sup> package for data wrangling and visualization, and the tidybayes (3.0.0)<sup>76</sup> and modelr (0.1.8)<sup>77</sup> packages for summarizing posterior distributions and predictions. We fit each model with two chains using within-chain parallelization, with 10,000 iterations each, 5,000 of which were for warmup, for a total of 10,000 samples post-warmup. Visual inspection showed that the chains were well-mixed, and for every parameter the effective number of samples in the bulk and tail of the distributions was >2,000, indicating effective sampling of the posterior.

**Transparency and ethics.** The study design and data analysis plan were not preregistered. Materials, de-identified data, R scripts, and statistical outputs are publicly available at <https://osf.io/eqftk/>. Study protocols were approved by the IRB at the University of Pennsylvania and permission to conduct research in Tanzania was approved by the Tanzanian Commission for Science and Technology. All methods were performed in accordance with the guidelines and ethical standards set out by relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. Permission to work in a camp was granted by each camp as a group and verbal informed consent to participate in research from each participant was obtained individually.

Received: 11 January 2022; Accepted: 15 April 2022

Published online: 24 May 2022

## References

1. Croley, J. A., Reese, V. & Wagner, R. F. Dermatologic features of classic movie villains. *JAMA Dermatol.* **153**, 559 (2017).
2. Dion, K., Berscheid, E. & Walster, E. What is beautiful is good. *J. Pers. Soc. Psychol.* **24**, 285–290 (1972).
3. Eagly, A. H., Ashmore, R. D., Makhijani, M. G. & Longo, L. C. What is beautiful is good, but ...: A meta-analytic review of research on the physical attractiveness stereotype. *Psychol. Bull.* **110**, 109–128 (1991).
4. Maestripieri, D., Henry, A. & Nickels, N. Explaining financial and prosocial biases in favor of attractive people: Interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. *Behav. Brain Sci.* **40**, e19 (2017).
5. Klebl, C., Rhee, J. J., Greenaway, K. H., Luo, Y. & Bastian, B. Beauty goes down to the core: Attractiveness biases moral character attributions. *J. Nonverbal Behav.* **46**, 83–97 (2022).
6. Workman, C. I. *et al.* Morality is in the eye of the beholder: The neurocognitive basis of the “anomalous-is-bad” stereotype. *Ann. N. Y. Acad. Sci.* **1494**, 3–17 (2021).
7. Hartung, F. *et al.* Behavioural and neural responses to facial disfigurement. *Sci. Rep.* **9**, 8021 (2019).
8. Jamrozik, A., Oraa, A. M., Sarwer, D. B. & Chatterjee, A. More than skin deep: Judgments of individuals with facial disfigurement. *Psychol. Aesthet. Creat. Arts* **13**, 117–129 (2019).
9. Rhodes, G. The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* **57**, 199–226 (2006).
10. Mengelkoch, S., Gassen, J., Prokosch, M. L., Boehm, G. W. & Hill, S. E. More than just a pretty face? The relationship between immune function and perceived facial attractiveness. *Proc. R. Soc. B Biol. Sci.* **289**, 11058 (2022).
11. Little, A. C., Apicella, C. L. & Marlowe, F. W. Preferences for symmetry in human faces in two cultures: Data from the UK and the Hadza, an isolated group of hunter-gatherers. *Proc. R. Soc. B Biol. Sci.* **274**, 3113–3117 (2007).
12. Apicella, C. L., Little, A. C. & Marlowe, F. W. Facial averageness and attractiveness in an isolated population of hunter-gatherers. *Perception* **36**, 1813–1820 (2007).
13. Tybur, J. M. *et al.* Re-evaluating the relationship between pathogen avoidance and preferences for facial symmetry and sexual dimorphism: A registered report. *Evol. Hum. Behav.* <https://doi.org/10.1016/j.evolhumbehav.2022.01.003> (2022).
14. Schaller, M. & Park, J. H. The behavioral immune system (and why it matters). *Curr. Dir. Psychol. Sci.* **20**, 99–103 (2011).
15. Neuberg, S. L., Kenrick, D. T. & Schaller, M. Human threat management systems: Self-protection and disease avoidance. *Neurosci. Biobehav. Rev.* **35**, 1042–1051 (2011).



16. Haselton, M. G., Nettle, D. & Murray, D. R. The Evolution of Cognitive Bias. in *The Handbook of Evolutionary Psychology* 1–20 (Wiley, 2015). <https://doi.org/10.1002/9781119125563.evpsych241>.
17. Haselton, M. G. & Nettle, D. The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personal. Soc. Psychol. Rev.* **10**, 47–66 (2006).
18. Zebrowitz, L. A. & Montepare, J. M. Social psychological face perception: Why appearance matters. *Soc. Personal. Psychol. Compass* **2**, 1497–1517 (2008).
19. Ryan, S., Oaten, M., Stevenson, R. J. & Case, T. I. Facial disfigurement is treated like an infectious disease. *Evol. Hum. Behav.* **33**, 639–646 (2012).
20. Shanmugarajah, K., Gaiind, S., Clarke, A. & Butler, P. E. M. The role of disgust emotions in the observer response to facial disfigurement. *Body Image* **9**, 455–461 (2012).
21. Klebl, C., Greenaway, K. H., Rhee, J. J. & Bastian, B. Ugliness Judgments Alert us to Cues of Pathogen Presence. *Soc. Psychol. Personal. Sci.* <https://doi.org/10.1177/1948550620931655> (2020).
22. Tybur, J. M., Lieberman, D. & Griskevicius, V. Microbes, mating, and morality: Individual differences in three functional domains of disgust. *J. Pers. Soc. Psychol.* **97**, 103–122 (2009).
23. Tybur, J. M., Lieberman, D., Kurzban, R. & DeScioli, P. Disgust: Evolved function and structure. *Psychol. Rev.* **120**, 65–84 (2013).
24. Navarrete, C. D. & Fessler, D. Meat is good to Taboo: Dietary proscriptions as a product of the interaction of psychological mechanisms and social processes. *J. Cogn. Cult.* **3**, 1–40 (2003).
25. Danovitch, J. & Bloom, P. Children's extension of disgust to physical and moral events. *Emotion* **9**, 107–112 (2009).
26. Haidt, J. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. (Vintage Books, 2012).
27. Krebs, D. *The Origins of Morality* (Oxford University Press, 2011). <https://doi.org/10.1093/acprof:oso/9780199778232.001.0001>.
28. Sanfey, A. G. The neural basis of economic decision-making in the Ultimatum Game. *Science (80-)* **300**, 1755–1758 (2003).
29. Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F. & Grafman, J. The neural basis of human moral cognition. *Nat. Rev. Neurosci.* **6**, 799–809 (2005).
30. May, J., Workman, C. I., Han, H., Haas, J. & Han, H. The Neuroscience of Moral Judgment: Empirical and Philosophical Developments. in *Neuroscience and Philosophy* (eds Sinnott-Armstrong, W. & De Brigard, F.) 17–47 (MIT Press, 2022). <https://doi.org/10.31234/osf.io/89jcx>.
31. Chapman, H. A., Kim, D. A., Susskind, J. M. & Anderson, A. K. In bad taste: Evidence for the oral origins of moral disgust. *Science (80-)* **323**, 1222–1226 (2009).
32. Cannon, P. R., Schnall, S. & White, M. Transgressions and expressions. *Soc. Psychol. Personal. Sci.* **2**, 325–331 (2011).
33. Molho, C., Tybur, J. M., Güler, E., Balliet, D. & Hofmann, W. Disgust and anger relate to different aggressive responses to moral violations. *Psychol. Sci.* **28**, 609–619 (2017).
34. Giner-Sorolla, R. & Chapman, H. A. Beyond purity: Moral disgust toward bad character. *Psychol. Sci.* **28**, 80–91 (2017).
35. Giner-Sorolla, R., Kupfer, T. & Sabo, J. What makes moral disgust special? An integrative functional review. *Adv. Exp. Soc. Psychol.* **57**, 223–289 (2018).
36. Apicella, C. L. & Barrett, H. C. Cross-cultural evolutionary psychology. *Curr. Opin. Psychol.* **7**, 92–97 (2016).
37. Henrich, J., Heine, S. J. & Norenzayan, A. The WEIRD people in the world? *Behav. Brain Sci.* **33**, 61–83 (2010).
38. Apicella, C., Norenzayan, A. & Henrich, J. Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evol. Hum. Behav.* **41**, 319–329 (2020).
39. Jaeger, B., Evans, A. M., Stel, M. & Beest, I. van. Who judges a book by its cover? Belief in physiognomy and reliance on trustworthiness impressions from faces. *PsyArXiv Prepr.* 1–19 (2019). <https://doi.org/10.31234/osf.io/8dq4x>.
40. Suzuki, A., Tsukamoto, S. & Takahashi, Y. Faces tell everything in a just and biologically determined world: Lay theories behind face reading. *Soc. Psychol. Personal. Sci.* **10**, 62–72 (2019).
41. Jones, B. C. *et al.* To which world regions does the valence–dominance model of social perception apply? *Nat. Hum. Behav.* **5**, 159–169 (2021).
42. Marlowe, F. Why the Hadza are Still Hunter-Gatherers. in *Ethnicity, Hunter-Gatherers, and the “Other”: Association or Assimilation in Africa* (ed. Kent, S.) 247–275 (Smithsonian Institution Press, 2002).
43. Weisz, J. R. East African medical attitudes. *Soc. Sci. Med.* **6**, 323–333 (1972).
44. Marlowe, F. W. W. *The Hadza: Hunter-Gatherers of Tanzania*. (University of California Press, 2010).
45. Ayeni, O. A., Ayeni, O. O. & Jackson, R. Observations on the procedural aspects and health effects of scarification in sub-Saharan Africa. *J. Cutan. Med. Surg.* **11**, 217–221 (2007).
46. Singh, D. & Bronstad, M. P. Sex differences in the anatomical locations of human body scarification and tattooing as a function of pathogen prevalence. *Evol. Hum. Behav.* **18**, 403–416 (1997).
47. Smith, K. M. & Apicella, C. L. Partner choice in human evolution: The role of cooperation, foraging ability, and culture in Hadza campmate preferences. *Evol. Hum. Behav.* **41**, 354–366 (2020).
48. Smith, K. M. & Apicella, C. L. Hadza hunter-gatherers disagree on perceptions of moral character. *Soc. Psychol. Personal. Sci.* **11**, 616–625 (2020).
49. Apicella, C. L., Azevedo, E. M., Christakis, N. A. & Fowler, J. H. Evolutionary origins of the endowment effect: Evidence from hunter-gatherers. *Am. Econ. Rev.* **104**, 1793–1805 (2014).
50. Apicella, C. L. High levels of rule-bending in a minimally religious and largely egalitarian forager population. *Religion. Brain Behav.* **8**, 133–148 (2018).
51. Stagnaro, M. N., Stibbard-Hawkes, D. N. E. & Apicella, C. L. Do religious and market-based institutions promote cooperation in Hadza hunter-gatherers? *Relig. Brain Behav.* (2022).
52. Tovee, M., Swami, V., Furnham, A. & Mangalparsad, R. Changing perceptions of attractiveness as observers are exposed to a different culture. *Evol. Hum. Behav.* **27**, 443–456 (2006).
53. Bissell, K. L. & Chung, J. Y. Americanized beauty? Predictors of perceived attractiveness from US and South Korean participants based on media exposure, ethnicity, and socio-cultural attitudes toward ideal beauty. *Asian J. Commun.* **19**, 227–247 (2009).
54. Marlowe, F. W. Hunter-gatherers and human evolution. *Evol. Anthropol. Issues News Rev.* **14**, 54–67 (2005).
55. Cushman, F. Deconstructing intent to reconstruct morality. *Curr. Opin. Psychol.* **6**, 97–103 (2015).
56. Hartman, R., Blakey, W. & Gray, K. Deconstructing moral character judgments. *Curr. Opin. Psychol.* **43**, 205–212 (2022).
57. Helzer, E. G. *et al.* Agreement on the perception of moral character. *Personal. Soc. Psychol. Bull.* **40**, 1698–1710 (2014).
58. Curtin, C. M. *et al.* Kinship intensity and the use of mental states in moral judgment across societies. *Evol. Hum. Behav.* **41**, 415–429 (2020).
59. Barrett, H. C. *et al.* Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proc. Natl. Acad. Sci.* **113**, 4688–4693 (2016).
60. Smith, K. M. & Apicella, C. L. Hadza hunter-gatherers are not deontologists and do not prefer deontologists as social partners. *J. Exp. Soc. Psychol.* **101**, 104314 (2022).
61. Gray, P. B. & Marlowe, F. Fluctuating asymmetry of a foraging population: The Hadza of Tanzania. *Ann. Hum. Biol.* **29**, 495–501 (2002).
62. Jones, N. B. *Demography and Evolutionary Ecology of Hadza Hunter-Gatherers*. (Cambridge University Press, 2016). <https://doi.org/10.1017/CBO9781107707030>.
63. Sznycer, D. & Patrick, C. The origins of criminal law. *Nat. Hum. Behav.* **4**, 506–516 (2020).

64. Hill, K. R., Wood, B. M., Baggio, J., Hurtado, A. M. & Boyd, R. T. Hunter-gatherer inter-band interaction rates: Implications for cumulative culture. *PLoS ONE* **9**, e102806 (2014).
65. Smith, K. M., Larroucau, T., Mabulla, I. A. & Apicella, C. L. Hunter-gatherers maintain assortativity in cooperation despite high levels of residential change and mixing. *Curr. Biol.* **28**, 3152–3157.e4 (2018).
66. Workman, C. I. & Chatterjee, A. The face image meta-database (fIMDb) & ChatLab facial anomaly database (CFAD): Tools for research on face perception and social stigma. *Methods Psychol.* **5**, 100063 (2021).
67. Kramer, R. S. S., Jenkins, R. & Burton, A. M. InterFace: A software package for face image warping, averaging, and principal components analysis. *Behav. Res. Methods* **49**, 2002–2011 (2017).
68. Tiddeman, B., Burt, M. & Perrett, D. Prototyping and transforming facial textures for perception research. *IEEE Comput. Graph. Appl.* **21**, 42–50 (2001).
69. McElreath, R. *Statistical Rethinking. Statistical Rethinking* (Chapman and Hall/CRC, 2018). <https://doi.org/10.1201/9781315372495>.
70. Makowski, D., Ben-Shachar, M. S., Chen, S. H. A. & Lüdtke, D. Indices of effect existence and significance in the Bayesian framework. *Front. Psychol.* **10**, 1125 (2019).
71. Kruschke, J. K. & Liddell, T. M. The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychon. Bull. Rev.* **25**, 178–206 (2018).
72. Kruschke, J. K. Rejecting or accepting parameter values in Bayesian estimation. *Adv. Methods Pract. Psychol. Sci.* **1**, 270–280 (2018).
73. Bürkner, P.-C. brms: An R package for Bayesian multilevel models using stan. *J. Stat. Softw.* **80**, 1–28 (2017).
74. Gabry, J. & Cesnovar, R. cmdstanr: R Interface to 'CmdStan'. (2020).
75. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
76. Kay, M. tidybayes: Tidy data and geoms for Bayesian models. *R package version 3.0.0* (2021) <https://doi.org/10.5281/zenodo.1308151>.
77. Wickham, H. modelr: Modelling Functions that Work with the Pipe. (2020).

## Acknowledgements

Research reported in this publication was supported by the National Institute of Dental & Craniofacial Research of the National Institutes of Health (F32DE029407 awarded to CIW), the Edwin and Fannie Gray Hall Center for Human Appearance at the University of Pennsylvania, and the University of Pennsylvania MindCORE program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders. We would like to thank Endeko Endeko, Victoria Maghali, and Ibrahim Mabulla for help in collecting data, and Audax Mabulla for providing administrative support in Tanzania.

## Author contributions

C.I.W. contributed to securing funding, conceiving and designing the study, creating the stimuli, and writing the manuscript, K.M.S. contributed to securing funding, conceiving and designing the study, data collection, data analysis, and writing the manuscript, C.L.A. contributed to conceiving and designing the study and writing the manuscript, A.C. contributed to securing funding, conceiving and designing the study, and writing the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to C.I.W. or K.M.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022