Journal of Articles in Support of the Null Hypothesis
Vol. 12, No. 1
 1539-8714
www.jasnh.com

Do People's First Names Match Their Faces?

Robin S. S. Kramer
University of York, UK
Alex L. Jones
Gettysburg College, PA, USA

We often feel that people's first names suit their faces in some way. Evidence has already shown that we share common stereotypes about how people with particular names should look. Here, we investigate whether there is any accuracy to these beliefs. Simply, can we match people's names to their faces? Across two experiments, we tested whether American (Experiment 1) and British participants (Experiment 2) were able to match the first names of strangers with photographs of their faces. Although Experiment 1 provided some initial support for accuracy in female participants, we were unable to replicate this result in Experiment 2. Therefore, we find no overall evidence to suggest that particular names and faces are associated with each other.

Key Words: first names, facial appearance, facial stereotypes, face perception

Corresponding author:
Dr. Robin Kramer, Department of Psychology, University of York, York YO10 5DD, UK
Email: remarknibor@gmail.com

"What's in a name?" Although Shakespeare might argue how unimportant or arbitrary names are, there is a growing body of research investigating the link between our names and who we are. Here, we focus on first names rather than surnames since the former demonstrate much more freedom and variety when it comes to selection while the latter tend to be determined more by lineage. Although evidence now suggests that people's names may influence their lives (e.g., Pelham, Carvallo, & Jones, 2005), few studies have explored the potential link between names and faces. The aim of the present study is to investigate whether there is any truth behind the common anecdote that people actually look like their names.

Researchers have found that names affect people's lives. For example, women with masculine names find more success in the legal profession (Coffey & McLaughlin, 2009). Boys with female-sounding names misbehave more at school (Figlio, 2007). Women with attractive first names may be seen as more physically attractive (Erwin, 1993; Garwood, Cox, Kaplan, Wasserman, & Sulzer, 1980). A name's ethnicity (e.g., Greg versus Jamal) may also have detrimental effects when applying for jobs (Bertrand & Mullainathan, 2004; cf. Fryer & Levitt, 2004). Our names may even play a role in where we choose to live or which jobs we pursue (Pelham, Mirenberg, & Jones, 2002).

If our first names can influence our lives, through how we perceive ourselves or how others perceive us, then it is important to consider how our names are chosen. Rather than arbitrarily, there are clear patterns regarding names and their owners. Names are generally associated with a specific sex (although a relatively small number are gender-neutral, for example, Jamie). In addition, parents' race and education influence name choices (Lieberson & Bell, 1992). Names also convey age information, as trends in naming fluctuate through the decades. For instance, the name Mildred suggests an older White woman in today's society. Gallagher and Chen (2008) provided a computer model with information from the US Social Security baby name database, which contains the 1000 most popular male and female baby names during the period 1880-2006. Simply by combining this information with the perceived age and gender of Americans in photographs, the model was able to predict first names with an overall accuracy of 82%. Our names may also represent our heritage, with first names being passed down through generations of families (Finch, 2008).

While names provide information regarding general category memberships (e.g., young, female, Black), do they also suggest more specific information about our facial appearance? We know that surnames were often derived from occupations (e.g., McCord & Shapter, 1964), but recently, researchers have shown that we also share facial stereotypes of occupation groups (Oldmeadow, Sutherland, & Young, 2012). Simply, people seem to agree on how the faces of those with certain occupations should look. More important for the current work, there is also evidence that we share facial stereotypes for first names (Lea, Thomas, Lamkin, & Bell, 2007). Participants were asked to construct faces (using computer software) for male first names. Those faces judged as good examples (by a second group) were averaged together to create a prototype for each name. These prototypes were subsequently matched with the correct names by a third group at above-chance levels. As such, we appear to have a shared idea of what people should look like based on their first names.

Computational approaches have demonstrated the utility of first names as facial attributes. By downloading photographs from Flickr, tagged with their owners' first names, computer models were able to use commonalities among images of people with the same

name to predict the names of new images at a level four times greater than chance (Chen, Gallagher, & Girod, 2014). Although this level of accuracy clearly benefited from naming conventions regarding sex, age, and ethnicity, it may well have also incorporated more specific commonalities arising from face–name associations.

In the present study, we investigate whether there is an association between faces and names above and beyond sex, age, and ethnicity. While often noted in real-world experiences, there is no experimental evidence attempting to answer this question to our knowledge.

Experiment 1

We utilised a set of Americans as models, and recruited a sample of American students as participants. It was important to make sure models and participants were of the same nationality, and also a similar age range, in order that they would be most experienced with any possible name-face relationships within their culture and age group if they existed.

Method

Participants

Fifty-seven undergraduate students at Gettysburg College volunteered to participate in exchange for either $10 or course credits. Of these, 52 self-reported as White Americans (age M = 19.92 years, SD = 1.22 years; 22 men), and their responses were included in subsequent analyses.

Stimuli

Two hundred and sixteen images were downloaded from an online database (www.facity.com) that contains high quality photographs of faces from cities around the world, along with their owners' first names and years of birth. All photographs in the database are posed front-on and with a neutral expression. Minimal or no cosmetics are worn. Images were cropped just below the chin and to the middle of the forehead. No clothing was visible.

Of these 216 images (all those in the database who lived in US cities), 86 individuals (43 men) were chosen as models because they were White and their birth years fell between 1980 and 1997. The website does not report when each photograph was taken, although we know this ranges from January 2010 (when the database extended beyond Germany) to April 2015 (when the images were downloaded). Therefore, the minimum and maximum ages of the models depicted in our photographs were 12 and 35 respectively. Due to this uncertainty regarding models' precise ages, our attempts to remove age as a potential cue to a person's name were only partially successful, given the maximum age range of 23 years.

Procedure

Participants were shown pairs of images on a laptop computer using custom PsychoPy software (Peirce, 2007). Models were paired at random for each participant with

a restriction that for each pairing, both models were of the same sex. In addition, the two models' names could not share the first three letters. This prevented pairs of models appearing together where both had the same name or the same root to their names.

On each of the 42 trials, two faces appeared onscreen along with one of the models' first names (chosen at random from the two models displayed). Participants were instructed to select which person's name it was ("Which of these individuals is named X?"). Responses were made using the mouse, and no time limit was given. The experiment took less than five minutes to complete.

Results and Discussion

Accuracy was calculated for each participant as the proportion of correct responses, separately for male trials (where pairs of men were shown) and female trials. These values were entered into a 2 (Participant Sex: Male, Female) x 2 (Model Sex: Male, Female) ANOVA. Participant Sex varied between subjects while Model Sex varied within subjects. We found a significant main effect of Participant Sex, $F(1, 50) = 6.96$, $p = .011$, $\eta^2_p = 0.12$, with women ($M = 0.537$, 95% CI [0.509, 0.565]) achieving higher accuracies than men ($M = 0.481$ [0.448, 0.513]). We found no main effect of Model Sex, $F(1, 50) = 1.00$, $p = .323$, $\eta^2_p = 0.02$, and no interaction between these two factors, $F(1, 50) = 0.61$, $p = .439$, $\eta^2_p = 0.01$.

Therefore, accuracies were aggregated across Model Sex to produce a combined proportion correct for all trials. One-sample t-tests were then carried out in order to determine whether women and men (separately) performed differently from chance (proportion correct of 0.5). We found that men performed at chance, $t(21) = 1.19$, $p = .246$, $d = 0.25$. However, women were significantly more accurate than what would be expected by chance, $t(29) = 2.66$, $p = .013$, $d = 0.49$. Even after applying Bonferroni correction for multiple tests (using $\alpha = .025$), this remained significant. It is interesting that women appear better able to accurately perceive cues to names from faces, given previous research demonstrating that women tend to outperform men when identifying facial expressions (Hall, 1978; Nowicki & Hartigan, 1988; Thayer & Johnsen, 2000).

Although the above result with female participants may be considered to have a medium effect size, we remain cautious in our interpretation. Given that previous research has shown that people share beliefs regarding how those with certain names should look (Lea et al., 2007), might it be possible that this could explain the significant accuracy found here? For example, if our female participants had a shared belief about which face goes with the name Bob, and in our sample of images, we happened to have a Bob that fit the description, then responses on this trial would tend to be correct. By chance, we might get such a name–face pairing in our set of stimuli that coincides with a common stereotype. As such, women would have a baseline score of one trial correct, and so performing at chance for the remaining trials would give an average proportion correct of 21.5 out of 42, or 0.512. Simply, a single model may drive their 'above chance' accuracy. If we suppose that chance performance were indeed 21.5/42, then we find that women do not differ from what we would expect by chance, $t(29) = 1.81$, $p = .080$, $d = 0.33$.

However, if we focus on the model as the unit of analysis, we fail to find support for this explanation. We calculated the proportion of correct trials for each model (where that model and their name appeared onscreen, paired with a second model), across all female participants' responses. Tests for normality, examining standardised skewness and

the Shapiro-Wilk test, indicated the data were statistically normal. In addition, we found no evidence of statistical outliers associated with high response accuracies. Therefore, there is no evidence to suggest that one or a few models represented exceptions, associated with particularly high proportions correct, and caused the above result. Indeed, similar analyses lead to the same conclusions regarding the male participants' responses, suggesting that their chance-level performance was not the result of one or a few low-scoring outliers bringing down an otherwise medium level of accuracy.

Another explanation for the present result may be that female participants were able to use age information in order to improve response accuracy. Given the possibility of an age difference of up to 23 years between the pair of models appearing onscreen, it may be that the name was (accurately) perceived to suit one model more because of their apparent age. We therefore analysed all female participants' trials, recording the age difference (which may not reflect the exact difference in ages at the times the photographs were taken – see above) between the two models on each trial and whether or not the response was correct. We carried out an independent samples t-test comparing correct and incorrect trials. However, we found no significant difference in the models' age gaps between correct (M = 4.73, 95% CI [4.47, 4.98]) and incorrect trials (M = 4.59, [4.31, 4.87]), $t(1258) = 0.70$, $p = .484$, $d = 0.04$. Therefore, although this does not rule out the possibility that age cues were useful, we find no supporting evidence for this explanation. Similar analyses also suggest that male participants made no use of age cues, $t(922) = 0.39$, $p = .698$, $d = 0.02$.

In order to increase our confidence that this finding with female participants is not simply a 'false positive', we decided to replicate the experiment (rather than relying solely on null hypothesis testing; Cohen, 1994) using a new set of models and participants. If the women again demonstrated above-chance accuracy, we could be more confident that observers are able to read cues to names in people's faces.

Experiment 2

We sought to replicate the results of our first experiment using a new set of British models, and a group of British participants as observers. Again, models and participants were of the same nationality and were of similar ages. We utilised a more restricted age range for our models in this second experiment in order to better control for any potential age-related cues to names. If we find accuracy in responses for a second time, this would both increase our confidence in the result and also demonstrate its generalisability across cultures.

Method

Participants

Forty-one students at the University of York volunteered to participate in exchange for chocolate biscuits. Of these, 40 self-reported as White British (age M = 22.63 years, SD = 3.94 years; 20 men), and their responses were included in subsequent analyses.

Stimuli

Ninety-six facial photographs were taken of undergraduate students at Bangor

University. Of these, 89 White British individuals (27 men; age range 19-24 years; age unavailable for 7 individuals) were used as stimuli. Before photographs were taken, individuals were asked to remove any cosmetics they were wearing, and were instructed to pose with a neutral expression, facing front-on to the camera. Images were cropped just below the chin, but the whole head, including hairstyle, was left visible. Minimal clothing around the shoulders also remained. First names were recorded. Students were awarded course credits for taking part.

Procedure

The same procedure was used as in Experiment 1, except that there were 44 trials here. For each participant, images were randomly paired, leaving one male photo unused (as there was an odd number of men in the stimulus set). As above, responses were made using the mouse, and no time limit was given. The experiment took less than five minutes to complete.

Results and Discussion

Accuracy was calculated for each participant as the proportion of correct responses, separately for male trials (where pairs of men were shown) and female trials. These values were entered into a 2 (Participant Sex: Male, Female) × 2 (Model Sex: Male, Female) ANOVA. Participant Sex varied between subjects while Model Sex varied within subjects.

We found no main effect of Participant Sex, $F(1, 38) = 0.46$, $p = .502$, $\eta^2_p = 0.01$, no main effect of Model Sex, $F(1, 38) = 2.80$, $p = .102$, $\eta^2_p = 0.07$, and no interaction between these two factors, $F(1, 38) = 2.05$, $p = .160$, $\eta^2_p = 0.05$. Therefore, accuracies were aggregated across Model Sex and Participant Sex to produce a combined proportion correct for all trials. A one-sample t-test was then carried out in order to determine whether participants (combined) performed differently from chance (proportion correct of 0.5). We found that our participants ($M = 0.500$, 95% CI [0.474, 0.526]) performed at chance, $t(39) = 0.00$. As such, we failed to replicate our first experiment, and found no evidence that participants were able to read cues to names from models' faces.

General Discussion

We know that first names provide information regarding numerous social categories, including sex, age, and ethnicity. Here, we investigated the relationship between first names and facial appearance. While previous research has shown that people share stereotypes regarding how the faces of those with particular names should look (Lea et al., 2007), we failed to extend this idea to show that those beliefs about name–face relationships are accurate.

In Experiment 1, we found evidence that American women may be able to identify American faces from first names at above-chance levels of accuracy. However, Experiment 2 failed to replicate this result with a new set of British faces and participants. While one could argue that such a result may therefore be culture-specific, we see no reason to draw such a conclusion. Instead, we believe the more likely explanation is simply that our initial significant effect was a false positive, and this is why we were unable to replicate.

Acknowledging the numerous anecdotes on this topic, why might we think there

could be an actual relationship between facial appearance and first names? First, it could be that facial features are statistically associated with different names. Even controlling for sex and ethnicity, names are passed down within families (i.e., Robert I, Robert II, etc.) and this could lead to specific names becoming associated with certain facial features (see Lea et al., 2007). Second, the characteristics of our names may lead others to perceive or treat us in certain ways ("Lilith is uptight"), and this may result in us behaving accordingly, including the ways in which we dress, groom, or express ourselves. Third, the reverse may be true and our faces may influence our names. If we look particularly childlike, for instance, we (or others) may choose to abbreviate our names to increase congruency (e.g., Timmy rather than Timothy). While one or more of these hypotheses may hold some truth, their effects are not large enough to produce accuracy in the current experiments.

In contrast, we can hypothesise many reasons why no (detectable) relationship exists between first names and facial appearance. For instance, there are many thousands of possible first names available for even a specific gender, age group, and ethnicity, and there are virtually unlimited facial morphologies. Even if these two categories were somehow related, it seems unlikely that we could develop any workable knowledge of the countless possible relationships, given the limited number we each could conceivably experience. Alternatively, there may simply be no relationship between faces and first names because parents select names based on personal preference, regardless of the appearance of the baby (which, of itself, may provide little insight into its adult appearance).

The lack of a relationship between first names and faces demonstrated here is particularly interesting given previous evidence that people hold shared stereotypes regarding how individuals with specific names should look (Lea et al., 2007). Our results suggest these stereotypes arise for reasons other than a genuine link between names and appearance. For example, cross-modal interaction between a name's sound and a person's facial features may lead us to picture a man called Bob (acoustically round-sounding) as round-faced (see Melara & Marks, 1990). The current findings may also have more practical relevance. Increasingly, sophisticated algorithms are employed to identify photographs posted to social networking sites, and their developers will utilise any available statistical patterns in order to improve such algorithms. Beyond simple category membership, we suggest first names will provide little use in these types of domains.

In conclusion, we find no overall evidence suggesting that particular first names and faces are associated, once sex, age, and ethnicity have been taken into account. So next time you meet someone who looks like a Bob, bear in mind that his name could just as easily be Steve.

References

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review, 94*, 991-1013.

Chen, H., Gallagher, A. C., & Girod, B. (2014). The hidden sides of names: Face modeling with first name attributes. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 36*, 1860-1873.

Coffey, B., & McLaughlin, P. A. (2009). Do masculine names help female lawyers become judges? Evidence from South Carolina. *American Law and Economics Review, 11*, 112-133.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49*, 997-1003.

Erwin, P. G. (1993). First names and perceptions of physical attractiveness. *The Journal of Psychology: Interdisciplinary and Applied, 127*, 625-631.

Figlio, D. N. (2007). Boys named Sue: Disruptive children and their peers. *Education Finance and Policy, 2*, 376-394.

Finch, J. (2008). Naming names: Kinship, individuality and personal names. *Sociology, 42*, 709-725.

Fryer, R. G., Jr., & Levitt, S. D. (2004). The causes and consequences of distinctively Black names. *The Quarterly Journal of Economics, 119*, 767-805.

Gallagher, A. C., & Chen, T. (2008). Estimating age, gender, and identity using first name priors. *Proceedings of the IEEE Computer Society: International Conference on Computer Vision and Pattern Recognition*. (Anchorage, AK, USA).

Garwood, S. G., Cox, L., Kaplan, V., Wasserman, N., & Sulzer, J. L. (1980). Beauty is only "name" deep: The effect of first-name on ratings of physical attraction. *Journal of Applied Social Psychology, 10,* 431-435.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin, 85*, 845-857.

Lea, M. A., Thomas, R. D., Lamkin, N. A., & Bell, A. (2007). Who do you look like? Evidence of facial stereotypes for male names. *Psychonomic Bulletin & Review, 14*, 901-907.

Lieberson, S., & Bell, E. O. (1992). Children's first names: An empirical study of social taste. *American Journal of Sociology, 98*, 511-554.

McCord, C. P., & Shapter, C. (1964). Family names derived from occupations: The origin of 250 English family names from crafts. *Archives of Environmental Health: An International Journal, 10,* 947-949.

Melara, R. D., & Marks, L. E. (1990). Processes underlying dimensional interactions: Correspondences between linguistic and nonlinguistic dimensions. *Memory & Cognition, 18,* 477-495.

Nowicki, S., Jr., & Hartigan, M. (1988). Accuracy of facial affect recognition as a function of locus of control orientation and anticipated interpersonal interaction. *The Journal of Social Psychology, 128*, 363-372.

Oldmeadow, J. A., Sutherland, C. A. M., & Young, A. W. (2012). Facial stereotype visualization through image averaging. *Social Psychological and Personality Science, 4,* 615-623.

Pelham, B. W., Carvallo, M., & Jones, J. T. (2005). Implicit egotism. *Current Directions in Psychological Science, 14*, 106-110.

Pelham, B. W., Mirenberg, M. C., & Jones, J. T. (2002). Why Susie sells seashells by the seashore: Implicit egotism and major life decisions. *Journal of Personality and Social Psychology, 82*, 469-487.

Peirce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods, 162*, 8-13.

Thayer, J., & Johnsen, B. H. (2000). Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology, 41*, 243-246.

Received: 5.31.2015
Revised: 6.14.2015
Accepted: 6.15.2015