Cognition 170 (2018) 1–8

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

What makes a face photo a 'good likeness'?

Kay L. Ritchie[a,b], Robin S.S. Kramer[a,b], A. Mike Burton[a,*]

[a] Department of Psychology, University of York, UK
[b] School of Psychology, University of Lincoln, UK

ARTICLE INFO

Keywords:
Facial likeness
Familiarity
Iconic images
Face averages
Prototype
Mental representation

ABSTRACT

Photographs of people are commonly said to be 'good likenesses' or 'poor likenesses', and this is a concept that we readily understand. Despite this, there has been no systematic investigation of what makes an image a good likeness, or of which cognitive processes are involved in making such a judgement. In three experiments, we investigate likeness judgements for different types of images: natural images of film stars (Experiment 1), images of film stars from specific films (Experiment 2), and iconic images and face averages (Experiment 3). In all three experiments, participants rated images for likeness and completed speeded name verification tasks. We consistently show that participants are faster to identify images which they have previously rated as a good likeness compared to a poor likeness. We also consistently show that the more familiar we are with someone, the higher likeness rating we give to *all* images of them. A key finding is that our perception of likeness is idiosyncratic (Experiments 1 and 2), and can be tied to our specific experience of each individual (Experiment 2). We argue that likeness judgements require a comparison between the stimulus and our own representation of the person, and that this representation differs according to our prior experience with that individual. This has theoretical implications for our understanding of how we represent familiar people, and practical implications for how we go about selecting images for identity purposes such as photo-ID.

1. Introduction

We all understand what it is to say that a particular image of someone is a good likeness. In fact, likeness is considered important for official forms of photo-ID, with passport-issuing offices around the world requiring someone familiar with the applicant to verify the likeness of the passport image (Australian Passport Office, 2012; HM Passport Office, 2014; Passport Canada, 2013). Despite this, there has been no systematic investigation of why observers pick out one image as a good likeness while considering another to be a bad likeness.

A number of different techniques in research on face recognition have used the concept of likeness as a key measure in the success of their manipulation. For example, studies manipulating the distinctiveness of face images (Allen, Brady, & Tredoux, 2009; Lee & Perrett, 2000), research on the caricaturing effect (Benson & Perrett, 1991; Rhodes, Brennan, & Carey, 1987), and research on face composites (Bruce, Ness, Hancock, Newman, & Rarity, 2002; Frowd et al., 2014) all used ratings of likeness as the dependent measure. Yet none have defined what this term means, relying on the fact that we all, participants and readers alike, intuitively understand the concept of likeness. Such an understanding seems to rely on the notion that a good likeness closely matches a canonical representation of a known person, or is perhaps some kind of 'super-stimulus' providing efficient access to such a representation, as has sometimes been claimed for caricatures.

When we consider different types of photos, we can see that one person can look very different across images depending on what they are trying to achieve in each (e.g., Leikas, Verkasalo, & Lönnqvist, 2013). For example, someone's passport photo will look different from their work website photo or their holiday photos. Previous work has shown that observers familiar with the person pictured can easily see that multiple, varied photos show the same person, whereas unfamiliar observers struggle to identify the same person across multiple images (e.g., Jenkins, White, van Montfort, & Burton, 2011; Laurence & Mondloch, 2016). The same is true even when only two images are pictured side by side and observers are asked to indicate whether the images show the same person or two different people. This task is easy for familiar viewers but significantly more difficult for unfamiliar viewers (Bruce, Henderson, Newman, & Burton, 2001; Clutterbuck & Johnston, 2002, 2004; Ritchie et al., 2015). Here, the difference between familiar and unfamiliar viewers seems to be their capacity to cope with a range of variability across different photos of the same person. With increased familiarity comes an ability to recognise a person from an increased range of images. Here we test whether familiarity also leads to an increased tolerance to the range of images one would categorise as being a good likeness.

* Corresponding author at: Department of Psychology, University of York, York YO10 5DD, UK.
E-mail address: mike.burton@york.ac.uk (A.M. Burton).

http://dx.doi.org/10.1016/j.cognition.2017.09.001
Received 21 April 2017; Received in revised form 31 August 2017; Accepted 4 September 2017

In an earlier systematic exploration of our ability to recognise and provide information (such as occupation) about familiar people from face images, the most common reason participants gave for failing to recognise a familiar person (71% of recognition failures) was that the photo was a 'bad likeness' (Hay, Young, & Ellis, 1991). The authors go on to quote participants' responses, giving examples such as "that's not what I remember him looking like", and "she's much younger in that photograph" (Hay et al., 1991, p. 778). One of the most interesting observations made was that different participants gave bad likeness responses to different images. This suggests that there is not something inherently poor about any given image, leading to all participants failing to recognise the same images. Rather, when each individual tries to recognise a familiar person, they compare the image under consideration with their own representation of that person. Each individual's representation of a person differs due to different levels of exposure to, or familiarity with, each target person, and so different images are a poor likeness for different observers.

Perhaps the most familiar face for a person is their own, and a recent study has shown that the images which participants select as a good likeness of themselves are, in fact, not optimal for identifying them (White, Burton, & Kemp, 2016). Participants ranked images of their own face for likeness, and a separate group of participants who were previously unfamiliar with these people also ranked the images for likeness after seeing a short video clip of each person. The images selected by this group as a good likeness, along with those selected by the participants themselves, were then used in a face matching task where two images were presented side by side, and a new group of participants were asked to judge whether or not the two images showed the same person. The images chosen by the participants themselves yielded lower accuracy on the matching task than the images selected by (unfamiliar) others. The results show that likeness judgements change as familiarity with the person pictured changes.

Jenkins et al. (2011) investigated the relationship between familiarity with celebrities and likeness ratings for multiple images of those celebrities. Participants rated multiple images of celebrities for likeness on a seven-point scale (extremely poor to extremely good likeness). Mean likeness ratings for images of each celebrity were positively correlated with the percentage of participants who recognised each celebrity. The results also showed that the variability in likeness ratings for multiple images of one celebrity could be greater than the variability in ratings between celebrities. Importantly, this approach used group means across participants. However, following from the earlier finding by Hay et al. (1991) that participants rated different images as a poor likeness, it may be more beneficial to analyse likeness ratings for each participant individually, rather than using group means.

What does it mean, therefore, for an image to be 'a good likeness'? Likeness judgements rely on a comparison between a physical stimulus and our own representation of an identity. The nature of this representation is unclear. As we become familiar with someone, we see multiple, variable instances of them which we need to be able to reconcile as the same person. Previous research has shown that when presented with multiple images of a person, viewers automatically extract the mean of the set (Kramer, Ritchie, & Burton, 2015), and it has been suggested that these averages provide us with stable representations of people (Burton, Jenkins, Hancock, & White, 2005; Frowd et al., 2014; Robertson, Kramer, & Burton, 2015). We might also represent some familiar celebrities through an 'iconic' or famous image (Allen et al., 2009; Carbon, 2008). In an experiment using common and uncommon images of celebrity faces, participants were able to recognise and name celebrities from commonly-seen or iconic images (~80% accuracy), but recognition rate dropped dramatically (~25%) when participants were shown uncommon images of the celebrities (Carbon, 2008). The results suggest that we recognise some celebrities from specific, commonly-seen images.

In the current research, we have sought to explore factors underlying likeness judgements for familiar faces. Underpinning this work is the idea that an image which is rated as being a good likeness should more closely resemble the rater's idea of what that person truly looks like in comparison with an image which is rated as a poorer likeness. We build on previous research which has shown that observers recognise a familiar individual across many variable images, suggesting that each of these images resembles their idea of what this person truly looks like. This study also follows on from the finding of Jenkins et al. (2011) that the more participants in a group who were familiar with a given celebrity, the higher the group likeness ratings for multiple images of that celebrity. These premises lead to three testable predictions: (1) it will be easier for an observer to recognise someone from an image which they perceive to be a good likeness of that person; (2) raters who are highly familiar with a person will have seen many images of them, and so will give higher likeness ratings to a larger range of images than will a less familiar viewer; (3) the specific images that each rater gives high likeness ratings to will be linked to their own experience of each person, and so likeness ratings will be different from one observer to the next.

In Experiment 1, we addressed the relationship between familiarity with a celebrity and likeness ratings given to images of them, as well as the idiosyncratic nature of likeness ratings. We examined individual and group variance in ratings (Hönekopp, 2006) to investigate the idiosyncratic nature of the perception of likeness. In addition, participants completed a speeded name verification task. If the perception of likeness from images maps on to the mental representation of identities, then it is reasonable to hypothesise that the higher the likeness rating, the faster that image will be verified as picturing the named identity. Experiment 2 explored the association between likeness ratings and observers' own prior experiences with each celebrity using images from films which participants had or had not seen. Finally, Experiment 3 looked at other types of images which have previously been suggested as candidates for representing familiar people (face averages) and certain celebrities (iconic images), and tested whether these types of images are given higher likeness ratings than other images.

2. Experiment 1 – The idiosyncratic nature of likeness perception

In order to test whether the perception of likeness is specific to each individual observer, we used a technique which allows us to differentiate between individual and group variance in likeness ratings. We asked observers to rate images for likeness twice, allowing us to determine whether variance in likeness ratings is explained predominantly by private or shared variance (following Hönekopp, 2006). We hypothesised that private perception of likeness (each person's idiosyncratic perceptions) would explain more variance than shared perceptions of likeness (agreement across the group) since what constitutes a good likeness could be different for each observer. If the perception of likeness is tied to each observer's familiarity with the person in question, observers may give higher likeness ratings for images of highly familiar compared with less familiar celebrities. We also tested whether likeness ratings were related to ease of recognition using a speeded name verification task. After rating the images for likeness, participates were shown the name of a celebrity followed by either an image of them or of another celebrity. If images rated as a good likeness are more easily recognised, we predicted that reaction times on the speeded name verification task would be faster for those images previously rated as a good compared to a poor likeness.

2.1. Method

2.1.1. Participants

Thirty-five participants (5 men; mean age: 19 years, range: 18–32 years) took part. All were students or other members of the University of York, UK, or the University of Lincoln, UK.