Research Article

THE ATTRACTIVENESS OF NONFACE AVERAGES: Implications for an Evolutionary Explanation of the Attractiveness of Average Faces

Jamin Halberstadt[1] and Gillian Rhodes[2]

[1]*University of Otago, Dunedin, New Zealand, and* [2]*University of Western Australia, Perth, Australia*

Abstract—*Researchers have argued that humans' attraction to average faces reflects an evolved psychological mechanism to identify high-quality mates. If this direct-selection account is correct, there is no reason to expect a similar averageness bias for stimuli that are irrelevant to reproductive fitness. The current study, however, found a strong relationship between averageness and attractiveness for dogs, wristwatches, and birds. The most parsimonious explanation is that humans have a general attraction to prototypical exemplars, and that their attraction to average faces is a reflection of this more general attraction. We tested whether a general preference for familiar stimuli can account for the attractiveness of averageness. This account was not supported for dogs or birds, but could not be ruled out for watches.*

Average facial configurations are attractive. Natural variations in averageness correlate significantly with attractiveness (Light, Hollander, & Kayra-Stuart, 1981; Rhodes, Sumich, & Byatt, 1999; Rhodes & Tremewan, 1996), digitally blended composite faces are more attractive than the individual faces from which they were created (Langlois & Roggman, 1990; Langlois, Roggman, & Musselman, 1994; Rhodes, Halberstadt, & Brajkovich, 2000), and faces can be made more (or less) attractive by increasing (or decreasing) their similarity to an average face (Rhodes & Tremewan, 1996; Rhodes et al., 1999).

Inevitably, researchers have considered the possibility that this preference for average faces has a biological basis (e.g., Langlois & Roggman, 1990; Thornhill & Gangestad, 1993). Because genetic and developmental abnormalities may be reflected in deviations from averageness, the argument goes, humans in the evolutionary past who were attracted to mates with extreme, asymmetric, or otherwise atypical faces would be at a reproductive disadvantage, relative to those attracted to average mates. Across many generations, the number of people desiring non-average-looking mates would consequently dwindle away.

The compelling logic of this account, however, obscures a vigorous debate over the evolution of psychological mechanisms (Barkow, Cosmides, & Tooby, 1992; Gould, 1978, 1991; Gould & Lewontin, 1979; Rose, Kamin, & Lewontin, 1984). A preference for averageness (and any other psychological mechanism) could have evolved in a variety of ways, other than because it was adaptive in humans' evolutionary past (see Endler & Basolo, 1998, for a recent review of such alternative accounts). Modeling studies confirm that preferences can evolve in the absence of any link between the preferred trait and mate quality (Johnstone, 1994). In these cases, the preference may emerge as a by-product of some more general feature of a recognition system, such as generalization gradients (Enquist & Johnstone, 1997), or the need to recognize potential mates (Johnstone, 1994) or objects seen from different viewpoints (Enquist & Arak, 1994).

For example, one possibility is that a preference for average or prototypical stimuli is a by-product of a preference for familiar stimuli. Perceivers do indeed judge (previously unseen) category prototypes as familiar (e.g., Franks & Bransford, 1971; Solso & McCarthy, 1981; Solso & Raynis, 1979), and it is well established that people respond positively to the familiar (the mere-exposure effect; Bornstein, 1989; Zajonc, 1968). Gordon and Holyoak (1983) reported findings consistent with this two-step argument: Exposure to unfamiliar stimuli (color matrices) increased liking both for the seen stimuli and for the unseen prototypes from which they were generated. Langlois et al. (1994) found that face composites were judged as more familiar than individual faces, and acknowledged that familiarity may very well explain the attractiveness of average faces. Furthermore, familiarity, which could inform a perceiver that a particular stimulus is not dangerous, is at least as evolutionarily plausible as averageness as a basis of facial attractiveness. However, if the attractiveness of averages is mediated by familiarity, then this has important implications for the evolutionary account outlined earlier. It suggests that what was selected in humans' evolutionary past was a preference for previously seen stimuli, rather than a preference for average faces per se, which would be only a by-product of this more general preference for the familiar.

Thus, there are at least two ways an attraction to average faces could have an evolutionary basis. First, the preference itself could have been directly selected, because averageness in faces carries information about mate quality. Second, the preference could be a by-product of some more general preference for averageness. We term these explanations the direct-selection and indirect-selection accounts, respectively. Unfortunately, because there is no way to trace the actual evolutionary history of facial attractiveness, we cannot compare these two accounts directly. What we can do is explore their logical implications: If attraction to average faces did indeed evolve as a solution to the problem of finding healthy mates, then there is no reason to expect average exemplars of other, evolutionarily irrelevant categories to be seen as attractive. However, if a preference for average faces reflects a more general feature of human information processing, then prototypical members of any category should be perceived as more attractive than unusual or extreme exemplars. The purpose of the present study was to test the plausibility of the direct-selection account by exploring the generality of people's attraction to average exemplars.

Surprisingly, the cognitive and social-cognitive literatures have little to say about the relationship between prototypicality and attractiveness. Some data come from studies of music preferences. Repp (1997, Study 2), for example, demonstrated that, as is the case with human faces, computer-averaged musical performances are judged

Address correspondence to Jamin Halberstadt, Department of Psychology, University of Otago, P.O. Box 56, Dunedin, New Zealand; e-mail: jhalbers@psy.otago.ac.nz.

Copyright © 2000 American Psychological Society

more positively than the individual performances used to create the averages. Smith and Melara (1990), however, found that the preference for musical prototypicality depended on expertise. Relative novices preferred prototypical to atypical chord progressions, although, interestingly, the reverse was true for experts (music graduate students).

Mandler's (1984) theory of emotion also predicts that somewhat atypical exemplars of a category should be preferred to the category prototype. According to this theory, evaluation is a function of both the match between a new stimulus and an existing schema and the cognitive activity required to resolve any incongruity. A mismatching stimulus produces arousal (by violating expectations and interfering with the completion of goals), which then takes on positive or negative valence depending on whether the new stimulus can be successfully integrated into an existing schema. A discrepant stimulus that can be assimilated into an existing schema will be responded to more positively than either a less discrepant stimulus that requires no cognitive integration or a more discrepant stimulus that cannot be fit to any existing prototype (Gaver & Mandler, 1987; see also Fiske, 1982).

Thus, the general relationship between averageness and attractiveness remains unclear. In the current studies, we examined this relationship using three very different stimulus categories: dogs, wristwatches, and birds. Dogs and watches were chosen as examples of familiar biological and artificial stimulus categories, respectively, whose exemplars vary consensually in both prototypicality and attractiveness. Birds were tested to address the possibility that attractiveness produces averageness, rather than the reverse. That is, in the case of dogs and watches (and people, for that matter), humans exert control over the frequency distribution of category members. If people are biased toward attractive exemplars, then these exemplars could become the central tendency of the category. In this account, people are attracted not to averageness per se, but rather to some configuration of features they find attractive for other reasons, and this configuration, by selective breeding or manufacturing, becomes average. Although the frequency of bird species is of course influenced by human behavior, there is no obvious way that this influence is associated with attractiveness.

If average faces are perceived as attractive because averageness provides information relevant to reproductive potential (the direct-selection account), there is no reason to expect a similar relationship between averageness and attractiveness to hold for other stimuli that are not relevant to reproductive success. If, however, a preference for averageness reflects some more general property of recognition systems (the indirect-selection account), then one might observe a positive correlation between averageness and attractiveness in any category for which a central tendency can be extracted. For faces, averageness has been judged on a variety of scales (e.g., "averageness," "typicality," "distinctiveness," "unusualness"), but in all cases, the judgments have correlated significantly (in the appropriate direction) with attractiveness (Light et al., 1981; Rhodes et al., 1999; Rhodes & Tremewan, 1996). These judgments have also been shown to change systematically with experimental manipulations of averageness, indicating the judgments' validity (Rhodes et al., 1999; Rhodes & Tremewan, 1996). In the present study, similar ratings were obtained for dogs, wristwatches, and birds.

Additionally, participants in the current study were asked to rate the subjective familiarity of each stimulus, so we could evaluate the possibility that average exemplars are attractive because they seem familiar. As we argued earlier, from the current perspective, the role of familiarity in averageness effects has critical implications for the direct-selection account. An explanation based only on familiarity suggests that the preference for average faces is a more general feature of the cognitive system, rather than an evolved strategy to take advantage of the genetic or developmental information average faces convey. The current studies, though correlational, could nevertheless rule out such an explanation if results showed that averageness accounted for significant variance in attractiveness ratings once the effect of familiarity had been partialed out. Furthermore, a correlational approach is appropriate because it is responses to natural variations in facial averageness that drive the evolution of any preference for facial averageness.

METHOD

Participants

Sixty-nine female and 15 male students at the University of Otago rated dogs and watches as part of their first-year psychology course. A separate sample of 42 female and 44 male Otago students was recruited from a student job clearinghouse on campus for the bird-rating procedure; these students were paid $10 for their participation in this and two other, unrelated studies.

Stimuli

Fifty drawings of dogs, 50 photos of wristwatches, and 50 drawings of birds, representing a range of attractiveness and typicality, were scanned at 72 dots per inch from three large reference books (Sylvester, 1993; Selby, 1994; and Perrins, 1990, respectively). Each dog was a different breed, drawn from a left-facing side or three-quarter perspective, and placed on a white 15.9-cm-square background. Watches were photographed from a full front perspective, and were placed on a white 10-cm-square background. Birds were drawn in either side or three-quarter perspective, unengaged in activity (such as feeding or mating), and were placed on a white 15.2-cm-square background. All taxonomic families of birds were represented in the stimulus set. All stimuli were surrounded by a 1-pixel black border.

Procedure

Participants were tested in private, light- and sound-attenuated cubicles. All stimuli and instructions were presented on Power Macintosh 7600 computers with 15-in. monitors, using custom-designed HyperCard software. Participants making dog and watch ratings were randomly assigned to judge both sets of stimuli on either attractiveness, averageness, or familiarity. Dogs were always rated before watches, but the order of presentation within each stimulus category was randomized for each participant. Participants made their judgments by clicking a computer mouse on 10-point "radio button" scales anchored at "very unattractive"–"very attractive," "very prototypical"–"very unusual" (later reverse-scored), or "very unfamiliar"–"very familiar." The procedure was identical for participants making bird ratings, except that only one set of stimuli was judged.

Prior to the rating task for each stimulus type, an instruction screen introduced the rating task, the use of the rating scale, and the meaning of the anchors. For example, participants rating averageness were told:

Table 1. *Descriptive statistics for the attractiveness, averageness, and familiarity ratings*

Measure	Stimulus set		
	Dogs	Watches	Birds
Attractiveness			
Mean (standard deviation)	5.2 (1.3)	4.6 (1.1)	5.5 (1.5)
Minimum, maximum	2.3, 7.8	2.3, 7.1	2.6, 8.4
Averageness			
Mean (standard deviation)	4.6 (1.7)	5.0 (1.0)	5.6 (2.0)
Minimum, maximum	2.1, 9.2	3.4, 7.4	1.2, 9.1
Familiarity			
Mean (standard deviation)	5.1 (1.7)	4.8 (2.1)	5.0 (2.6)
Minimum, maximum	1.4, 8.0	1.1, 8.8	1.3, 10.0

Note. Ratings were made on 10-point scales anchored at "very unattractive"–"very attractive," "very prototypical"–"very unusual" (reverse-scored), or "very unfamiliar"–"very familiar."

If someone asked you to think of a typical dog, what would that dog look like? This is your "average" or "prototypical" dog. Some dogs look relatively similar to this prototypical dog. Other dogs look relatively unusual or distinct from this prototypical dog. The computer will show you a series of dogs on the screen. Please judge how UNUSUAL-LOOKING you perceive each dog to be, using the following 10-point scale. Click on the point on the scale that corresponds to how unusual-looking each dog seems to you. The LESS unusual-looking (the more prototypical or average) the dog is, the farther to the LEFT you should click on the scale. The MORE unusual-looking the dog is, the farther to the RIGHT you should click on the scale.

Each picture was presented in the center of the screen above a prompt (e.g., "How attractive is this dog?") and the appropriate rating scale. Participants could change their rating for any stimulus as many times as they wished before clicking a button labeled "continue" to advance to the next stimulus, but they could not return to a previously rated item. Participants were encouraged to use the full range of the scale.

RESULTS

Interrater reliability was assessed separately for averageness, attractiveness, and familiarity judgments, using Cronbach's coefficient alpha. Reliability was generally high for all stimulus sets (.96, .92, and .94, respectively, for dogs; .77, .89, and .97 for watches; .96, .97, and .98 for birds). Therefore, for each rating scale, ratings were averaged across participants to get a mean rating for each stimulus. The mean, standard deviation, and range of each measure appear in Table 1. The zero-order and partial correlations among the measures appear in Table 2.

As seen in Table 2, dogs' averageness and familiarity (which were themselves positively correlated, $r[50] = .63$, $p < .001$) both strongly predicted their attractiveness. Furthermore, when the effect of familiarity was partialed out, the correlation between attractiveness and averageness remained significant. When the effect of averageness was partialed out, however, the correlation between attractiveness and familiarity was not significant.

These analyses were repeated on the watch data. As was the case with dogs, both averageness and familiarity correlated strongly with attractiveness (and with each other, $r[50] = .87$, $p < .001$). However, partial correlations indicated that it was familiarity, not averageness, that independently predicted attractiveness. With averageness statistically controlled, the correlation between attractiveness and familiarity remained significant, but with familiarity controlled, the correlation between attractiveness and averageness did not.

For birds, averageness was again strongly positively correlated with attractiveness, but, in contrast to the results for dogs and watches, the correlation between familiarity and attractiveness, though positive, did not reach significance, $p = .16$ (averageness and familiarity themselves correlated .41, $p < .005$). Furthermore, partial correlations revealed that averageness predicted attractiveness over and above familiarity. When the effects of averageness were statistically controlled, the correlation between familiarity and attractiveness literally dropped to zero.

DISCUSSION

The purpose of this study was to test the proposal that humans' demonstrated attraction to mathematically average faces evolved as a direct solution to the adaptive problem of identifying high-quality mates (e.g., Thornhill & Gangestad, 1993). We argued that if this

Table 2. *Zero-order (and partial) correlations with attractiveness*

Measure	Stimulus set		
	Dogs	Watches	Birds
Averageness	.69** (.58)**	.65** (.13)	.50** (.47)**
Familiarity	.47** (.06)	.69** (.34)*	.20 (.00)

Note. Familiarity was partialed out of the zero-order correlations between attractiveness and averageness; averageness was partialed out of the zero-order correlations between attractiveness and familiarity.
*$p < .05$. **$p < .001$.

direct-selection account were correct, the preference for averageness should be limited to human faces. Average dogs, watches, and birds, for example, should not necessarily be seen as attractive. The data, however, clearly show that they are. Three very different stimulus sets all revealed a positive relationship between averageness and attractiveness at least as strong as that found for human faces (e.g., Rhodes et al., 1999, reported zero-order correlations of .45 between averageness and attractiveness and −.52 between distinctiveness and attractiveness). Furthermore, the present study ruled out the possibility, at least in the case of birds, that human behavior (breeding practices, production standards, etc.) merely increased the frequency of some stimulus configurations that were attractive for other reasons.

It is of course possible that different mechanisms underlie the attractiveness effects in each of these stimulus domains and for human faces. Average faces could be preferred for their information value, whereas average dogs, watches, and birds could be attractive for other reasons. However, the most parsimonious account of the data is that humans are attracted to facial averageness not because it correlates with mate quality, but as a consequence of a more general perceptual or cognitive processing bias. The recent report in this journal that facial attractiveness does not correlate with health raises similar doubts about whether facial preferences are adaptive (Kalick, Zebrowitz, Langlois, & Johnson, 1998).

If people do not prefer average faces because they signal mate quality, what other reason could there be? One possibility, tested in the present experiments, is that a preference for averageness reflects a more general preference for familiar stimuli. Prototypical exemplars feel familiar, and the positive affect derived from familiarity might be generalized to physical-attractiveness judgments. Gordon and Holyoak (1983) found that the mere-exposure effect, the attachment of positive affect to previously seen stimuli, did in fact generalize to the unseen prototype from which the seen stimuli were generated. The current data, however, suggest that the averageness bias is not, in general, mediated by subjective familiarity. For both dogs and birds, averageness explained unique variance in attractiveness judgments when the effects of familiarity were partialed out, indicating that prototypical exemplars were not attractive merely because they seemed familiar. This account remains a possibility for watches, although it is not clear why familiarity played a more important role in this stimulus set. The difference may be due to the biological status of dogs and birds, or to the ease with which prototypes can be abstracted from these categories.

Our data do not rule out an account in which a preference for averageness reflects a bias toward stimuli resembling those that have been previously processed by the visual system, irrespective of their subjective familiarity. Kunst-Wilson and Zajonc (1980) have shown that people prefer previously seen polygons, even when their recognition of the polygons is at chance levels. We note, however, that in a recent study (Rhodes et al., 2000), we failed to find any evidence that the attractiveness of averaged composite faces was due to their similarity to seen faces.

If a preference for averageness does not reflect a general bias toward familiar stimuli, then what other general feature of information processing might account for it? Langlois and her colleagues have suggested that the preference for averageness can be explained by an innate tendency to form prototypes (e.g., Langlois & Roggman, 1990; Rubenstein, Kalakanis, & Langlois, 1999), although the question of just why such prototypes should be attractive remains open. Other possibilities are suggested by the modeling studies described in the introduction. Those studies showed that symmetry preferences could emerge as a by-product of the need to recognize objects from different viewpoints (Enquist & Arak, 1994), to recognize potential mates (Johnstone, 1994), or to generalize across similar exemplars (Enquist & Johnstone, 1997). Given that average configurations are often highly symmetric (except when directional asymmetries are present), similar accounts may help explain why average exemplars are attractive. We note, however, that in the case of faces, averageness and symmetry make unique contributions to attractiveness (Rhodes et al., 1999), suggesting that these accounts may not fully explain the attractiveness of averageness. Testing these conjectures is beyond the scope of the present report, but our results strongly suggest that future research should investigate alternatives to the adaptationist account of the attractiveness of average faces.

Acknowledgments—This research was supported by 1998 and 1999 Otago Research Grants and grants from the Division of Sciences and the Psychology Department at the University of Otago to the first author and by a grant from the Australian Research Council to the second author. We acknowledge Gary Bartlett and Cindy Hall for their assistance in all phases of this research, and Patricia Haden for her help in software design. We also thank Steven Catty, Melanie Hills, and Veronica Soto for their assistance with stimulus development and data collection, and Paula Niedenthal, who provided helpful comments on a draft of this article.

REFERENCES

Barkow, J.H., Cosmides, L., & Tooby, J. (1992). *The adapted mind.* New York: Oxford University Press.

Bornstein, R.F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Review, 106,* 265–289.

Endler, J.A., & Basolo, A.L. (1998). Sensory ecology, receiver biases and sexual selection. *Trends in Ecology & Evolution, 13,* 415–420.

Enquist, M., & Arak, A. (1994). Symmetry, beauty and evolution. *Nature, 372,* 169–172.

Enquist, M., & Johnstone, R.A. (1997). Generalization and the evolution of symmetry preferences. *Proceedings of the Royal Society of London B, 264,* 1345–1348.

Fiske, S.T. (1982). Schema-triggered affect: Applications to social perception. In M.S. Clark & S.T. Fiske (Eds.), *Affect and cognition: The 17th Annual Carnegie Symposium on Cognition* (pp. 55–78). Hillsdale, NJ: Erlbaum.

Franks, J.J., & Bransford, J.D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology, 90,* 65–74.

Gaver, W.W., & Mandler, G. (1987). Play it again, Sam: On liking music. *Cognition and Emotion, 1,* 259–282.

Gordon, P.C., & Holyoak, K.J. (1983). Implicit learning and generalization of the "mere exposure" effect. *Journal of Personality and Social Psychology, 45,* 492–500.

Gould, S.J. (1978). Sociobiology: The art of storytelling. *New Scientist, 16,* 530–533.

Gould, S.J. (1991). Exaptation: A crucial tool for an evolutionary psychology. *Journal of Social Issues, 47,* 43–65.

Gould, S.J., & Lewontin, R.C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London, 205,* 581–598.

Johnstone, R.A. (1994). Female preference for symmetrical males as a byproduct of selection for mate recognition. *Nature, 372,* 172–175.

Kalick, S.M., Zebrowitz, L.A., Langlois, J.H., & Johnson, R.M. (1998). Does human facial attractiveness honestly advertise health? Longitudinal data on an evolutionary question. *Psychological Science, 9,* 8–13.

Kunst-Wilson, W.R., & Zajonc, R.B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science, 207,* 557–558.

Langlois, J.H., & Roggman, L.A. (1990). Attractive faces are only average. *Psychological Science, 1,* 115–121.

Langlois, J.H., Roggman, L.A., & Musselman, L. (1994). What is average and what is not average about attractive faces? *Psychological Science, 5,* 214–220.

Light, L.L., Hollander, S., & Kayra-Stuart, F. (1981). Why attractive people are harder to remember. *Personality and Social Psychology Bulletin, 7,* 269–276.

Mandler, G. (1984). *Mind and body.* New York: W.W. Norton.

Perrins, C.M. (1990). *The illustrated encyclopedia of birds: The definitive guide to birds of the world*. London: Headline.

Repp, B.H. (1997). The aesthetic quality of a quantitatively average music performance: Two preliminary experiments. *Music Perception, 14*, 419–444.

Rhodes, G., Halberstadt, J.H., & Brajkovich, G. (2000). *Generalization of mere exposure effects in social stimuli*. Manuscript submitted for publication.

Rhodes, G., Sumich, A., & Byatt, G. (1999). Are average facial configurations attractive only because of their symmetry? *Psychological Science, 10*, 52–58.

Rhodes, G., & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science, 7*, 105–110.

Rose, S., Kamin, L.J., & Lewontin, R.C. (1984). *Not in our genes: Biology, ideology and human nature.* Harmondsworth, England: Penguin Books.

Rubenstein, A.J., Kalakanis, L., & Langlois, J.H. (1999). Infant preferences for attractive faces: A cognitive explanation. *Developmental Psychology, 15*, 848–855.

Selby, I. de L.S. (1994). *Wrist watches*. Philadelphia: Courage Books.

Smith, J.D., & Melara, R.J. (1990). Aesthetic preference and syntactic prototypicality in music: 'Tis the gift to be simple. *Cognition, 34*, 279–298.

Solso, R.L., & McCarthy, J.E. (1981). Prototype formation: Central tendency model vs. attribute frequency model. *Bulletin of the Psychonomic Society, 17*, 10–11.

Solso, R.L., & Raynis, S.A. (1979). Prototype formation from imaged, kinesthetically, and visually presented geometric figures. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 701–712.

Sylvester, P. (1993). *The Reader's Digest illustrated book of dogs* (2nd rev. ed.). Montreal, Canada: Reader's Digest.

Thornhill, R., & Gangestad, S.W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Human Nature, 4*, 237–269.

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*, 1–27.

(RECEIVED 6/8/99; REVISION ACCEPTED 11/11/99)