

It's not just average faces that are attractive: Computer-manipulated averageness makes birds, fish, and automobiles attractive

JAMIN HALBERSTADT

University of Otago, Dunedin, New Zealand

and

GILLIAN RHODES

University of Western Australia, Perth, Australia

Average faces are attractive. We sought to distinguish whether this preference is an adaptation for finding high-quality mates (the *direct selection* account) or whether it reflects more general information-processing mechanisms. In three experiments, we examined the attractiveness of birds, fish, and automobiles whose averageness had been manipulated using digital image manipulation techniques common in research on facial attractiveness. Both manipulated averageness and rated averageness were strongly associated with attractiveness in all three stimulus categories. In addition, for birds and fish, but not for automobiles, the correlation between subjective averageness and attractiveness remained significant when the effect of subjective familiarity was partialled out. The results suggest that at least two mechanisms contribute to the attractiveness of average exemplars. One is a general preference for familiar stimuli, which contributes to the appeal of averageness in all three categories. The other is a preference for averageness per se, which was found for birds and fish, but not for automobiles, and may reflect a preference for features signaling genetic quality in living organisms, including conspecifics.

The question of what makes human faces attractive has intrigued psychologists, philosophers, biologists, and artists alike. However, one answer, based on recent psychological research, is surprising: averageness. Natural variations in the averageness of facial configurations consistently correlate with attractiveness, both within and across cultures (Jones & Hill, 1993; Light, Hollander, & Kayra-Stuart, 1981; Rhodes, Harwood, Yoshikawa, Nishitani, & McLean, 2002; Rhodes, Sumich, & Byatt, 1999; Rhodes & Tremewan, 1996; Rhodes, Yoshikawa, et al., 2001; but see Grammer & Thornhill, 1994). Faces can be made more (or less) attractive by increasing (or decreasing) their similarity to a computer-averaged composite of multiple same-sex faces (Rhodes et al., 1999; Rhodes & Tremewan, 1996). Furthermore, the computer-averaged faces themselves are usually more attractive than the individual faces that compose them and become increasingly attractive as additional faces are added to the composite (Langlois & Roggman, 1990; Rhodes et al., 2002; Rhodes et al., 1999).¹

The appeal of the seemingly mundane trait of averageness is sobering for the romantic, but it is logical for

the evolutionary psychologist. From a functional perspective, averageness is a sign of developmental stability and genetic quality (Møller & Swaddle, 1997; Rhodes, Zebrowitz, et al., 2000), deviations from which may be reflected in asymmetries, distinctive features, or other departures from a prototypical face (e.g., Møller & Swaddle, 1997). If, in our evolutionary past, the congenital conditions that such nonprototypical features reflect were associated with reduced reproductive fitness, people preferring those features in mates would have been less likely to reproduce successfully. If, in turn, face preferences are heritable, a preference for averageness would have evolved over time (e.g., Langlois & Roggman, 1990).

Recently, Halberstadt and Rhodes (2000) challenged this account, which they termed the *direct selection hypothesis*. They reasoned that if attraction to average faces did indeed evolve as a solution to the problem of finding healthy mates, there is no reason to expect average exemplars of other, evolutionarily irrelevant categories to be seen as attractive. The researchers collected averageness and attractiveness ratings of a wide variety of dogs, birds, and wristwatches and found strong positive correlations (at least as strong as those for faces) in each category they studied. Because preferences for average dogs, birds, and watches confer no obvious reproductive benefits, the researchers argued that these averageness preferences may reflect some general underlying feature of human information processing, rather than a specific adaptation to the problem of mate choice.

This research was supported by an Otago Research Grant to the first author and by a grant from the Australian Research Council to the second author. We acknowledge Cindy Hall, Billy Jorgensen, Evonne Miller, and Steve Catty for their assistance in all phases of this research. Correspondence regarding this research should be directed to J. Halberstadt, Department of Psychology, University of Otago, P. O. Box 56, Dunedin, New Zealand (e-mail: jhalbers@psy.otago.ac.nz).

A critic might argue, however, that although Halberstadt and Rhodes's (2000) stimuli were more attractive when they were average, they were not necessarily more attractive *because* they were average. A third variable, such as functionality, could be the basis for both judgments. For example, dogs that cannot bark, birds that cannot fly, and watches that cannot tell time are not only unusual, but also do not "work" well. Alternatively, the implied causality in Halberstadt and Rhodes's results could be reversed: The central tendency of a category could be responsive to humans' judgments of attractiveness, as in the case of selective breeding of dogs (or indeed, humans). Therefore, the implications of Halberstadt and Rhodes's data for the direct selection hypothesis are limited, because it is implicit in the hypothesis that attractiveness judgments are actually caused by perceived averageness, the demonstration of which requires experimental control. Otherwise, it is plausible that attraction is responsive to (i.e., caused by) averageness only in faces, just as the direct selection hypothesis predicts.

The goal of the present experiments was to exert such control. In three experiments, we manipulated the averageness of stimuli in three nonface categories—birds, fish, and automobiles—via two computer-imaging techniques used to manipulate averageness in faces: (1) increasing and decreasing the similarity of a stimulus to a known average stimulus (e.g., Rhodes et al., 1999; Rhodes & Tremewan, 1996) and (2) combining increasingly large sets of stimuli to create averaged composites (Langlois & Roggman, 1990; Rhodes et al., 2002). An increase in attractiveness as a function of manipulated similarity to an average stimulus or of the number of faces used to create an averaged stimulus would provide clear evidence that humans respond to averageness in stimuli other than faces and would, at a minimum, show that the direct selection hypothesis must be an incomplete account of why averageness is attractive.

If there is a generic preference for averageness, what might account for it? One possibility is humans' preference for the familiar. We know that faces and other stimuli become more attractive when their familiarity is increased by repeated exposure (Bornstein, 1989; Rhodes, Halberstadt, & Brajkovich, 2001; Zajonc, 1968) and, also, that average exemplars and composites appear familiar (e.g., Franks & Bransford, 1971; Langlois, Roggman, & Musselman, 1994). Therefore, subjective familiarity could potentially account for the appeal of average exemplars. Consistent with this hypothesis, Halberstadt and Rhodes (2000) found that familiarity contributed to the appeal of average exemplars in all three classes examined (wristwatches, dogs, and birds). Furthermore, for wristwatches, the correlation between averageness and attractiveness disappeared when subjective familiarity was controlled, indicating that familiarity could entirely account for the attractiveness of average wristwatches. For dogs and birds, however, a significant correlation remained between averageness and attractiveness when subjective familiarity was controlled.

These results suggest the intriguing possibility that familiarity might play a different role in the attractiveness of averageness in animals and artifacts. For artifacts, the appeal of average exemplars might simply reflect the fact that manufacturers tend to produce items that people like (for whatever reason). In this case, a preference for average exemplars would not be a preference for averageness per se but, rather, for some configuration of features that is attractive for other reasons, which by selective manufacturing becomes average. For animals, however, the appeal of average exemplars cannot be explained in this way. A second goal of the present experiments, therefore, was to seek independent support for these conjectures by examining the contribution of subjective familiarity to the attractiveness of average birds, fish, and automobiles.

EXPERIMENT 1

One technique that has been used to manipulate the averageness of faces is to vary their similarity to an average face. Rhodes and Tremewan (1996), for example, used a computerized caricature generator to distort line drawings of faces toward or away from an average (i.e., a composite of all of the faces used in the study). Their program compared *landmark points* on each face with corresponding points on the composite and then either exaggerated the differences between the two (to create caricatures) or reduced them (to create anticaricatures; see Rhodes, Brennan, & Carey, 1987, for details). Rhodes and Tremewan's (1996) participants judged the anticaricatured faces as more attractive than the undistorted versions and the undistorted versions as more attractive than the caricatured versions.

In Experiment 1, the participants rated the attractiveness of caricatured and anticaricatured line drawings of passerines (the largest family of birds), created with the very same caricature generator as that described above. Seven versions of each bird were developed: an undistorted drawing of the bird, three caricatured drawings that exaggerated (by 17%, 33%, or 50%) the differences in shape between that passerine and an average passerine (a computer-averaged composite of the 14 passerines used in the study), and three anticaricatured drawings that reduced those differences (by 17%, 33%, or 50%). A separate set of participants judged the averageness of the stimuli, in order to assess the correlation between these judgments and the attractiveness ratings independently of our experimental manipulation of the stimuli.

In addition, a third set of participants judged the familiarity of all the stimuli, so that we could assess the contribution of subjective familiarity to any observed relationship between averageness and attractiveness in the stimuli. Halberstadt and Rhodes (2000) found that familiarity contributed to, but did not fully explain, the attractiveness of average birds. We expected to replicate that result here.

Method

Participants. One hundred and fifty-five male and female participants were recruited from a student job clearinghouse on the

Otago University campus. They received \$10 for taking part in this and several other, unrelated procedures, which together took approximately 1 h.

Stimuli. The stimuli consisted of 14 sets of line drawings of resident New Zealand passerines—blackbird, brown creeper, chaffinch, fantail, kokako, pipit, robin, rock wren, rook, saddleback, sparrow, stitchbird, tomtit, and tui—originally created by Rhodes and McLean (1990, Study 3). Each set consisted of seven stimuli: a veridical line drawing of a bird in profile, taken from one of several publications on New Zealand birds; 17%, 33%, and 50% caricatures of the bird; and 17%, 33%, and 50% anticaricatures of the bird. Caricatures were created by distorting the veridical drawing of a bird the corresponding proportions away from the average (a composite of the 14 undistorted line drawings); anticaricatures were created by distorting the same drawings toward the average. (See Rhodes & McLean, 1990, and Rhodes et al., 1987, for a complete description of the caricaturing and anticaricaturing procedures.)

Rhodes and McLean's (1990) original 98 stimuli, printed on paper, were scanned at 72 dots per inch and were placed on white backgrounds approximately 13 cm wide and 10 cm high, with a 1-pixel black border. Examples (reduced in size) appear in Table 1.

Procedure. The participants were tested in private, light- and sound-attenuated experimental cubicles. All stimuli and instructions were presented on Power Macintosh 7600 computers with 15-in. monitors, using custom-designed HyperCard software. The participants were randomly assigned to judge the stimuli, which were presented in a different random order for each person, on attractiveness, averageness, or familiarity ($n =$ approximately 50 in each judgment condition). The participants made their judgments on 10-point "radio button" scales anchored at *very unattractive/very attractive*, *very prototypical/very unusual* (reverse-scored), or *very unfamiliar/very familiar*.

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

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

turn to a previously rated item. The participants were encouraged to use the full range of the scale. The ratings took approximately 10 min.

Results and Discussion

Interrater reliability was assessed separately for averageness, attractiveness, and familiarity judgments, using Cronbach's coefficient alpha. Reliability was high for all three ratings scales (.96, .94, and .97, respectively), and the judgments were, therefore, averaged across participants to get mean ratings for each averageness level (50% caricature, 33% caricature, 17% caricature, veridical drawing, 17% anticaricature, 33% anticaricature, and 50% anticaricature) for each bird. These means were submitted to separate analyses of variance (ANOVAs), with distortion level as a repeated measures factor, which revealed highly significant effects on all three ratings [$F(6,294) = 102.03$ (averageness), $F(6,324) = 192.78$ (attractiveness), and $F(6,294) = 376.07$ (familiarity), $ps < .001$; see Table 1]. Paired t tests indicated that all means, in all rating tasks, differed at $p < .001$ (Bonferroni corrected significance level for multiple comparisons is .007).

When averageness, attractiveness, and familiarity ratings were correlated, averageness and familiarity [which were themselves positively correlated; $r(98) = .95$, $p < .001$] both strongly predicted attractiveness [$rs(98) = .94$ and $.93$, $ps < .001$]. Furthermore, when the effect of familiarity or of averageness was partialled out, the relationship between the remaining two variables weakened but remained significant [$rs(98) = .46$ and $.42$, $ps < .001$].

Experiment 1 therefore demonstrates that participants' attraction to passerines depends in a nearly linear way on their physical similarity to an average passerine. Because we controlled that similarity—indeed, with the same software as that used to manipulate averageness in faces—there can be no question about the causal relationship. Manipulated averageness produced differences in attractiveness, paralleling Rhodes and Tremewan's (1996) study of facial attractiveness and clearly showing that these effects are not limited to human faces. In addition, the correlational data indicated that, independent of our manipulation, averageness and attractiveness were strongly associated and, more important, that average-

Table 1
Mean Ratings (With Standard Deviations) of Attractiveness, Averageness, and Familiarity of Passerine Birds as a Function of Caricature Level in Experiment 1

Trait	Example (Caricature Level)													
	(+50%)		(+33%)		(+17%)		(0%)		(-17%)		(-33%)		(-50%)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Attractiveness	3.5	1.4	4.0	1.3	4.5	1.1	4.8	1.2	5.4	1.1	5.7	1.2	6.3	1.22
Averageness	2.7	1.1	3.6	1.3	4.2	1.3	5.0	1.2	5.7	1.0	6.3	1.1	7.2	1.00
Familiarity	3.5	1.1	4.2	1.1	4.5	1.2	5.2	1.1	5.8	1.1	6.5	1.1	7.3	1.20

ness predicted attractiveness over and above familiarity. These data provide the first independent confirmation of Halberstadt and Rhodes's (2000) conclusion that average stimuli are not attractive merely because they seem more familiar to the perceiver, although their familiarity does contribute to their appeal.

EXPERIMENT 2

Experiment 1 manipulated averageness by varying similarity to an averaged composite of all the stimuli. A second, and in some respects better, technique for manipulating averageness is to create digital composites of varying numbers of faces. Langlois and Roggman (1990), for example, created six degrees of averageness by combining sets of 2, 4, 8, 16, and 32 faces. Participants' attractiveness ratings were linearly related to the number of component faces in a composite, and the 16- and 32-face composites were significantly more attractive than the average rating of the individual faces used to create them (and more attractive than most of the individual faces themselves). In the remaining two experiments, we used a similar technique, systematically increasing the number of component exemplars in computer-averaged composites to manipulate the averageness of fish and automobiles. Thus, the participants in Experiment 2 judged the attractiveness of undistorted line drawings of fish, along with composites of 2, 4, 8, 16, and 32 fish. If averageness is attractive in fish, as it is in human faces, we would expect to find an increase in attractiveness as the number of component fish in the image increases. As in Experiment 1, separate participants rated the averageness and the familiarity of the fish, so that we could examine the contribution of subjective familiarity to the attractiveness of averageness.

Method

Participants. Seventy-three male and female University of Otago students participated as part of the assessment for their first-year psychology course. Four participants did not provide complete data and were not included in the following analyses.

Stimuli. Thirty-two color photographs of fish were scanned from two reference books (Last, Scott, & Talbot, 1983; Nelson, 1994). Between 150 and 200 landmark points were mapped to the body, fins, and gills of each fish and, using the digital blending software Morph (1994), were joined to create line drawings (examples appear in Figure 1). From these we created averaged composites containing 2 (Av2, $n = 16$), 4 (Av4, $n = 8$), 8 (Av8, $n = 4$), 16 (Av16, $n = 2$), and 32 (Av32, $n = 1$) component fish. Each Av2 fish was created by generating a new line drawing from the average location of the corresponding landmark points on each of 2 component fish (selected from the original 32 without replacement). Av4 fish were, in turn, created by morphing pairs of Av2 stimuli, and so on for the higher order composites. This process, the results of which are illustrated in Figure 1, is explained in more detail in Rhodes et al. (1999).

Procedure. The procedure was identical to that in Experiment 1. Once again, the participants were randomly assigned to rate the averageness, attractiveness, or familiarity of all 63 stimuli, presented in a different random order for each participant. The ratings took approximately 10 min and were gathered in conjunction with several other, unrelated procedures.

Results and Discussion

Interrater reliability was high for averageness, attractiveness, and familiarity judgments (alphas = .95, .96, and .95, respectively), and the ratings were, therefore, averaged across participants within each judgment and each generation of stimulus. The results, shown in Table 2, were analyzed in three separate ANOVAs, with image type (individual, Av2, Av4, Av8, Av16, and Av32) as a repeated measures factor, which revealed main effects of image type in all cases [$F(5,105) = 18.44$ (averageness), $F(5,115) = 11.07$ (attractiveness), and $F(5,110) = 8.05$ (familiarity), $ps < .001$]. As can be seen in the table, all three dependent measures increased across the first three generations of stimuli but generally leveled off subsequently. This asymptotic pattern is perhaps not surprising, because, with repeated averaging, the stimuli may become less distinct, particularly to the nonexpert eye (see Figure 1). It is possible that, had our original stimuli been more variable or our participants more expert, attractiveness, averageness, and familiarity would have continued to increase through to the Av32 composite. In any case, an asymptote effect is also typically found with faces, so these results are quite consistent with the face data (e.g., Langlois & Roggman, 1990; Rhodes et al., 2002).

Correlational analyses revealed that, as in Experiment 1 and Halberstadt and Rhodes's (2000) research, averageness and familiarity both predicted unique variance in attractiveness. Both were both positively correlated with attractiveness [$rs(63) = .75$ (averageness) and $.46$ (familiarity), $ps < .001$] even when the effects of the second variable were partialled out [$rs(63) = .69$ and $.26$, $p < .001$ and $p < .05$, respectively].

EXPERIMENT 3

The design of Experiment 3 was identical to that of Experiment 2, except that the stimuli were automobiles. Our aim was to determine whether attractiveness was affected by manipulated averageness even in this artificial class and to assess the contribution of subjective familiarity to the relationship between attractiveness and averageness.

Method

Participants. Eighty-nine male and female University of Otago students participated as part of the assessment for their first-year psychology course.

Stimuli. Sixty-three line drawings of automobiles, representing 19 different automobile manufacturers, were created using the techniques described in Experiment 2. First, 32 digital images were obtained by photographing automobiles on display at Dunedin second-hand dealerships ($n = 11$) and downloading photographs from second-hand dealers' and corporations' Web sites ($n = 14$ and 7). All the automobiles used were four-door sedans manufactured between 1990 and 2000, inclusive, and were photographed in full profile. As in Experiment 2, approximately 120 landmark points were mapped and joined to create line drawings of each automobile, which were then averaged to create higher order composite drawings (examples appear in Figure 2).

Procedure. The procedure, run in conjunction with several other, unrelated experiments, was identical to that in Experiment 2.

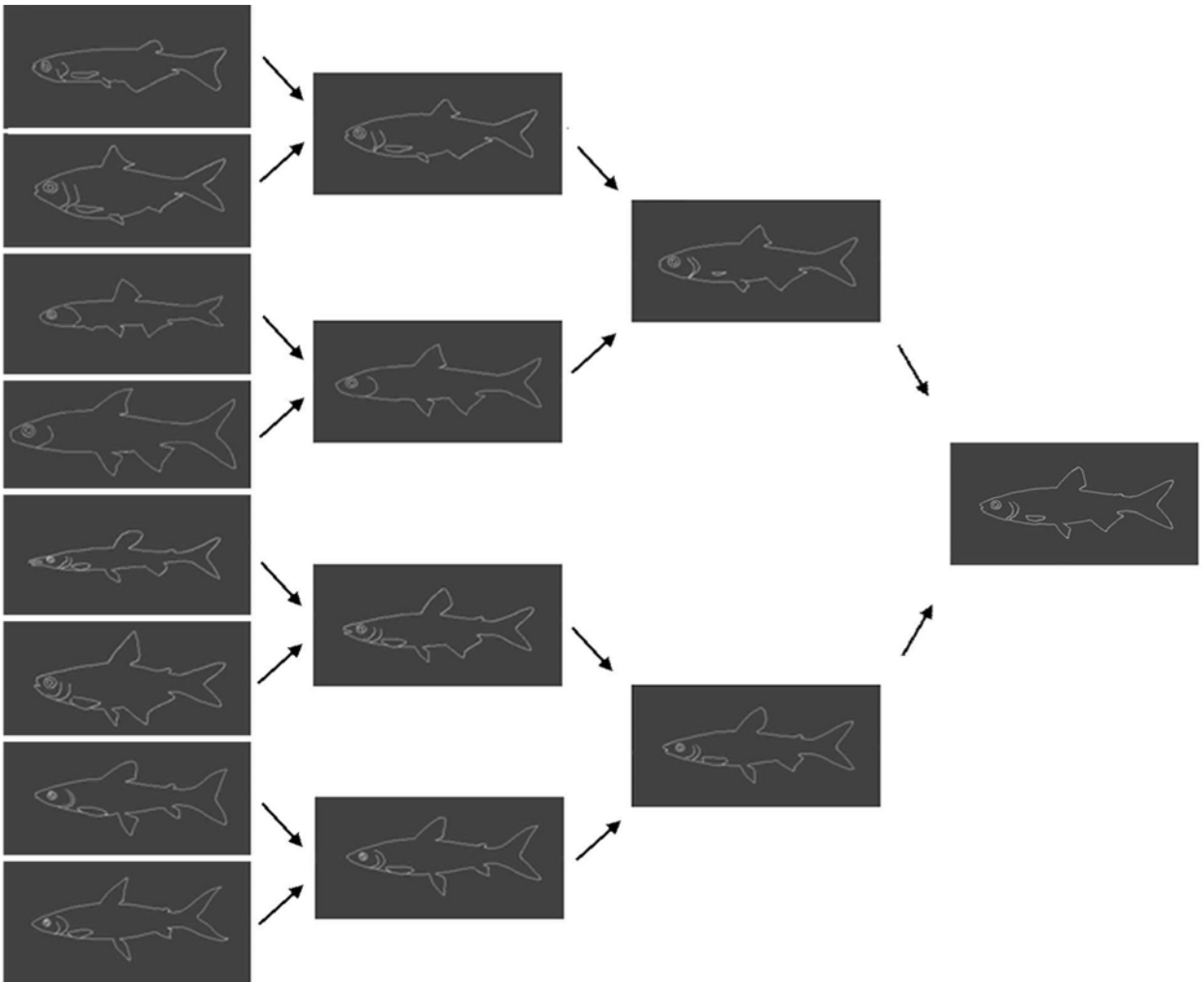


Figure 1. One branch of the fish family tree.

Results and Discussion

Averageness, attractiveness, and familiarity judgments once again showed high interjudge reliability (alphas = .96, .97, and .97, respectively) and were averaged across participants within each judgment and each type of stimulus. Separate within-subjects ANOVAs revealed main effects for all three measures [$F(5,120) = 7.09$ (averageness), $F(5,195) = 6.46$ (attractiveness), and $F(5,115) = 13.99$ (familiarity), $ps < .001$]. As can be seen in Table 2, there were significant differences on all three dependent measures between ratings of the original automobiles and ratings of the Av2 composites. Higher order composites generally did not differ from one another, as might be expected given the uniformity of the original stimuli (see Figure 2). These results are consistent both with the results of Experiments 1 and 2 and with ratings of face composites created with similar digital averaging techniques (e.g., Langlois & Roggman, 1990). Thus, even in a stim-

ulus class as arbitrary, artificial, and homogeneous as late-model, four-door sedans, a strong relationship emerged between manipulated averageness and attractiveness.

Correlational analyses revealed that averageness correlated strongly with attractiveness [$r(63) = .54$, $p < .001$], but in contrast to the first two experiments, that correlation disappeared when the effect of familiarity was partialled out [$r(63) = .006$]. On the other hand, familiarity correlated with attractiveness [$r(63) = .74$, $p < .001$] over and above the effect of averageness [partial $r(63) = .57$, $p < .001$].

GENERAL DISCUSSION

Halberstadt and Rhodes's (2000) discovery of averageness-attractiveness correlations in nonface categories was interpreted as evidence against the hypothesis that humans' preference for average faces is an adaptation for

Table 2
Mean Ratings (With Standard Deviations) of Attractiveness, Averageness, and Familiarity
as a Function of Image Type in Experiments 2 and 3

Trait	Image Type											
	Original Stimuli		Av2		Av4		Av8		Av16		Av32	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experiment 2 (Fish)												
Attractiveness	3.9 _a	1.1	4.4 _b	1.2	5.2 _c	1.3	5.3 _c	1.5	5.0 _c	1.5	5.1 _{bc}	1.9
Averageness	4.0 _a	1.2	5.2 _b	1.1	5.7 _c	1.1	6.2 _d	1.2	6.5 _d	1.5	6.1 _{bcd}	2.1
Familiarity	4.1 _a	1.0	5.1 _b	1.3	5.3 _b	1.5	5.7 _b	2.0	5.9 _b	2.1	5.7 _b	2.5
Experiment 3 (Autos)												
Attractiveness	4.3 _a	1.3	4.8 _b	1.3	4.8 _b	1.3	4.8 _b	1.4	4.9 _b	1.5	4.4 _a	1.8
Averageness	4.6 _a	1.4	6.1 _b	1.2	6.0 _b	1.3	5.9 _b	1.4	5.8 _b	1.9	5.9 _b	2.0
Familiarity	5.2 _a	1.4	6.2 _b	1.6	6.4 _{bc}	1.7	6.6 _{cd}	1.7	6.0 _{bc}	2.0	7.3 _d	2.1

Note—Means within a row that do not share a common subscript differ at $p < .01$.

identifying high-quality mates (the direct selection hypothesis). Halberstadt and Rhodes's data, however, could not rule out a number of alternative explanations that would make their reported correlations compatible with

direct selection. The present experiments do. When techniques used for increasing and decreasing the averageness of faces were applied to nonface stimuli, the effects on attractiveness paralleled those found for faces. In-

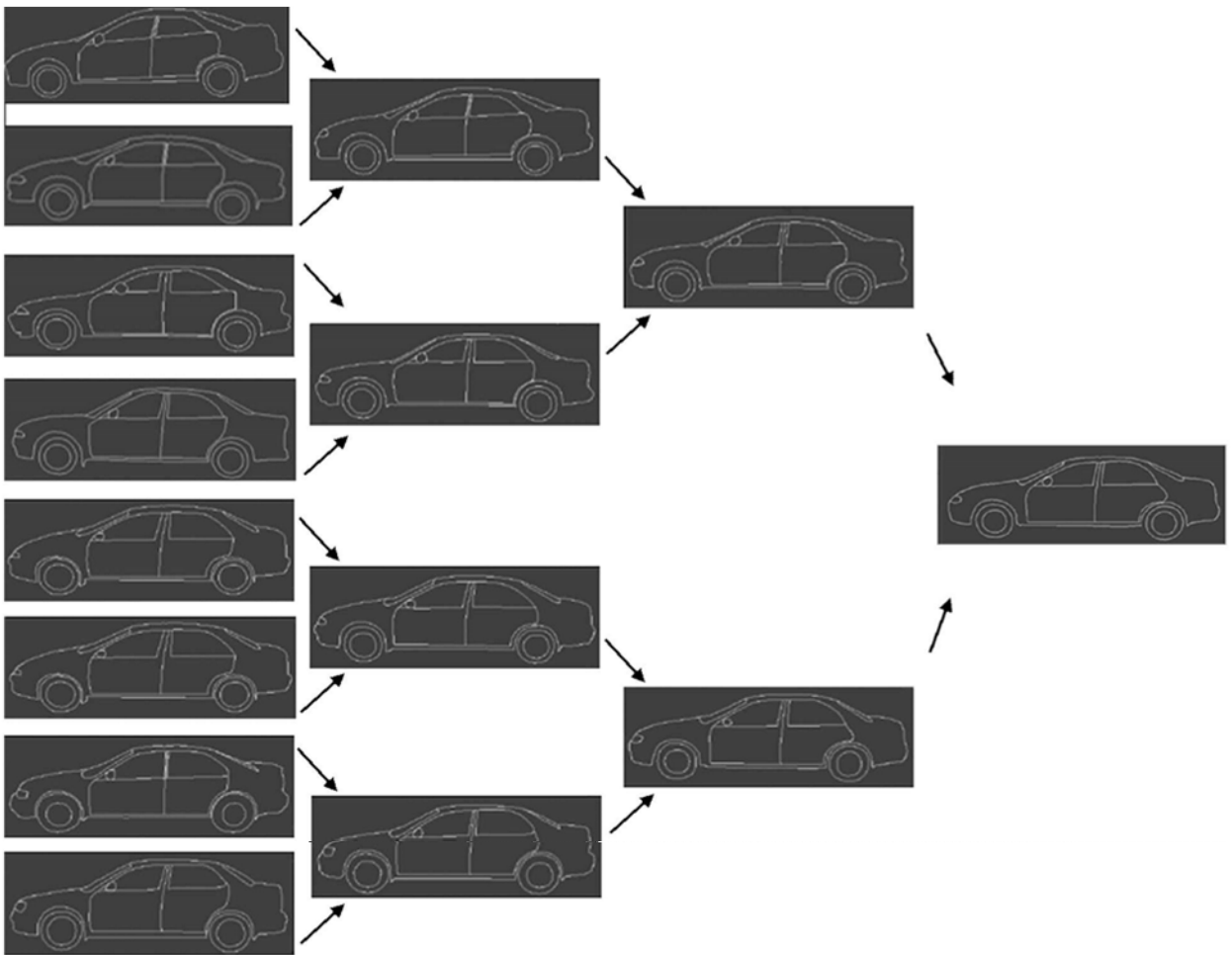


Figure 2. One branch of the automobile family tree.

creasing the averageness of passerines, fish, and automobiles, by increasing their similarity to an average stimulus or by adding more exemplars to a composite image, increased attractiveness; decreasing the averageness of these stimuli, by reducing their similarity to an average stimulus, reduced attractiveness. A parsimonious account of the attractiveness of averageness must now explain both face and nonface preferences.

A critic might argue that our image manipulation techniques themselves introduced alternative explanations for the effects. Similar criticisms, in fact, have been leveled at studies in which averageness in faces has been manipulated and have been part of an ongoing debate about the validity of those manipulations (e.g., Alley & Cunningham, 1991; Langlois et al., 1994). Alley and Cunningham, for example, argued, among other things, that the process of digitally blending produces emergent features, such as symmetry and smooth complexion, in the resulting composites, which could account for their attractiveness. Although our manipulations of averageness could be confounded by such features (although the use of line drawings eliminates some of them), this possibility is irrelevant to the goals of the experiments. Whatever accounts for the attractiveness of averaged faces—whether it be averageness per se or some emergent feature associated with the averaging process—the effect that averaging manipulations have on attractiveness is not specific to faces and, therefore, cannot be used as evidence to support a direct selection account.

The second goal of the present experiments was to provide an independent test of the role of familiarity in the relationship between averageness and attractiveness. This relationship remained significant when familiarity was partialled out of judgments of passerines and fish, but not of automobiles, analogous to the results of Halberstadt and Rhodes's (2000) analysis. Thus, familiarity can fully account for the attractiveness of both classes of artifacts examined to date (automobiles and wristwatches), but not of the three classes of animals examined (birds, fish, and dogs). The apparent close link between familiarity and attractiveness in artifacts may reflect the deliberate, selective manufacture of items with attractive traits. According to this view, neither averageness nor familiarity, per se, is what makes average artifacts attractive; rather, both are linked to the distribution of traits that is perceived as attractive for whatever evolutionary, sociological, or cultural reasons. However, given the strong appeal of familiarity (Bornstein, 1989; Zajonc, 1968) it seems likely that it would also contribute to the attractiveness of average artifacts. Either way, the results suggest that averageness per se does not account for their appeal.

On the other hand, in animals (including humans), a preference for averageness, which is well known to reflect health and developmental stability (e.g., Møller & Swaddle, 1997), could be a means of identifying organisms of high genetic quality. According to this account, a preference for averageness has general functional value.

In the case of conspecifics, it may indeed be a tool for finding high-quality mates, with healthy individuals providing direct benefits in parental care and, possibly, indirect genetic benefits to offspring, if their resistance to disease is heritable. In the case of other species, averageness might have helped people to avoid interaction with or consumption of diseased animals, to select quality livestock for breeding, or to assess the health and threat of predators. At this point, of course, this account is speculative, and we are continuing to study the differences between natural and artificial categories more generally.

To conclude, we suggest that at least two mechanisms may contribute to the attractiveness of average exemplars. The first is a generic preference for familiar stimuli, which plays a role in judgments of both artifacts and animals, including humans. This preference could be due either to the appeal of familiarity itself or to the relationship among familiarity, averageness, and manufacturing standards. The second mechanism, which is independent of familiarity, is associated with attractiveness judgments of animals, including humans, and may represent an evolved preference for signs of genetic and/or phenotypic quality. Regardless of whether these specific mechanisms are supported by further research, the present data indicate that the direct selection hypothesis is incomplete as an account of why we like average faces.

REFERENCES

- ALLEY, T. R., & CUNNINGHAM, M. R. (1991). Averaged faces are attractive, but very attractive faces are not average. *Psychological Science*, *2*, 123-125.
- BORNSTEIN, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, *106*, 265-289.
- CUNNINGHAM, M. R., BARBEE, A. P., & PIKE, C. L. (1990). What do women want? Facial metric assessment of multiple motives in the perception of male facial physical attractiveness. *Journal of Personality & Social Psychology*, *59*, 61-72.
- FRANKS, J. J., & BRANSFORD, S. D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology*, *90*, 65-74.
- GRAMMER, K., & THORNHILL, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology*, *108*, 233-242.
- HALBERSTADT, J., & RHODES, G. (2000). The attractiveness of non-face averages: Implications for an evolutionary explanation of the attractiveness of average faces. *Psychological Science*, *11*, 285-289.
- JOHNSTON, V. S., HAGEL, R., FRANKLIN, M., FINK, B., & GRAMMER, K. (2001). Male facial attractiveness: Evidence for hormone-mediated adaptive design. *Evolution & Human Behavior*, *22*, 251-267.
- JONES, D., & HILL, K. (1993). Criteria of facial attractiveness in five populations. *Human Nature*, *4*, 271-296.
- 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.
- LAST, P. R., SCOTT, E. O. G., & TALBOT, F. H. (1983). *Fishes of Tasmania*. Hobart, Tasmania: Tasmanian Fisheries Development Authority.
- LIGHT, L. L., HOLLANDER, S., & KAYRA-STUART, F. (1981). Why attractive people are harder to remember. *Personality & Social Psychology Bulletin*, *7*, 269-276.
- MØLLER, A. P., & SWADDLE, J. P. (1997). *Asymmetry, developmental stability and evolution*. Oxford: Oxford University Press.
- MORPH (Version 2.5) [Computer software] (1994). San Diego: Gryphon

Software.

- NELSON, J. S. (1994). *Fishes of the world* (3rd ed.). New York: Wiley.
- PENTON-VOAK, I. S., PERRETT, D. I., CASTLES, D. L., KOBAYASHI, T., BURT, D. M., MURRAY, L. K., & RINAMISAWA, R. (1999). Menstrual cycle alters face preference. *Nature*, **399**, 741-742.
- PERRETT, D. I., LEE, K. J., PENTON-VOAK, I., ROWLAND, D., YOSHIKAWA, S., BURT, D. M., HENZI, S. P., CASTLES, D., & AKAMATSU, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, **394**, 884-887.
- PERRETT, D. I., MAY, K. A., & YOSHIKAWA, S. (1994). Facial shape and judgments of female attractiveness. *Nature*, **368**, 239-242.
- RHODES, G., BRENNAN, S., & CAREY, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology*, **19**, 473-497.
- RHODES, G., HALBERSTADT, J., & BRAJKOVICH, G. (2001). Generalization of mere exposure effects in social stimuli. *Social Cognition*, **19**, 57-70.
- RHODES, G., HARWOOD, K., YOSHIKAWA, S., NISHITANI, M., & MCLEAN, I. G. (2002). The attractiveness of average facial configurations: Cross-cultural evidence and the biology of beauty. In G. Rhodes & L. A. Zebrowitz (Eds.), *Advances in visual cognition: Vol. 1. Facial attractiveness: Evolutionary, cognitive, and social perspectives* (pp. 35-58). Westport, CT: Ablex.
- RHODES, G., HICKFORD, C., & JEFFERY, L. (2000). Sex-typicality and attractiveness: Are supermale and superfemale faces super-attractive? *British Journal of Psychology*, **91**, 125-140.
- RHODES, G., & MCLEAN, I. G. (1990). Distinctiveness and expertise effects with homogeneous stimuli: Towards a model of configural coding. *Perception*, **19**, 773-794.
- RHODES, G., SUMICH, A., & BYATT, G. (1999). Are average facial configurations only attractive because of their symmetry? *Psychological Science*, **10**, 52-58.
- RHODES, G., & TREMEWAN, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science*, **7**, 105-110.
- RHODES, G., YOSHIKAWA, S., CLARK, A., LEE, K., MCKAY, R., & AKAMATSU, S. (2001). Attractiveness of facial averageness and symmetry in non-Western populations: In search of biologically based standards of beauty. *Perception*, **30**, 611-625.
- RHODES, G., ZEBROWITZ, L. A., CLARK, A., KALICK, S. M., HIGHTOWER, A., & MCKAY, R. (2000). Do facial averageness and symmetry signal health? *Evolution & Human Behavior*, **21**, 1-16.
- ZAJONC, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality & Social Psychology*, **9**, 1-27.

NOTE

1. Average faces are attractive, but they may not be optimally attractive. For example, female faces with extreme sex-typical traits are preferred (albeit slightly) over those with average traits (Perrett et al., 1998; Perrett, May, & Yoshikawa, 1994; Rhodes, Hickford, & Jeffery, 2000). Some studies have also found that extreme sex-typical traits are attractive in male faces (Cunningham, Barbee, & Pike, 1990; Johnston, Hagel, Franklin, Fink, & Grammer, 2001), but this is controversial, with other studies reporting a preference for feminized male faces (Penton-Voak et al., 1999; Perrett et al., 1998; Rhodes, Hickford, & Jeffery, 2000). Even if average faces are not optimally attractive, however, they are more attractive than most faces.

(Manuscript received July 20, 2001;
revision accepted for publication March 4, 2002.)