*Bulletin of the Psychonomic Society*
*1984, 22 (6), 549-552*

# Facial stereotypes of good guys and bad guys: A replication and extension

ALVIN G. GOLDSTEIN, JUNE E. CHANCE, and BARBARA GILBERT
*University of Missouri, Columbia, Missouri*

This research tests the hypothesis that faces can serve as stimuli triggering consensual stereotypical responses in observers. Facial photographs of white middle-aged men were arranged in five separate arrays, each containing equal numbers of portraits. Subjects were first led to believe that each array contained one and only one portrait of each of the following "occupations": mass murderer, armed robber, rapist, medical doctor, clergyman, and engineer. Following these instructions, subjects were asked to choose one picture for each of the six occupations. Chi-square analysis of their choices indicates that in all five arrays subjects' selections are significantly nonrandom. A small number of portraits are often selected as "bad guys," and a small number of other portraits are often selected as "good guys." In addition, not only are criminal and noncriminal face prototypes suggested by the data, but subjects' choices also tend to be occupation specific; however, this tendency was more evident when the choices were among the criminal rather than the noncriminal occupations. The results of the study were discussed in terms of the probability that stereotyping is a factor in jury decisions.

It is widely recognized that physical appearance is an important factor in interpersonal relations and social cognition. A large literature is devoted to the general topic of impression formation, and a portion of that literature includes several studies assessing the relationship between subjects' perceptions of others' facial characteristics and their attitudes or impressions about them (Fiske & Taylor, 1984; Hamilton, 1981; Schneider, Hastorf, & Ellsworth, 1979). In general, these studies suggest that, in the absence of other relevant information, people seem to use faces to make judgments about other individuals' personality, occupation, potential behavior, intelligence, and much more. In other words, to paraphrase Strachey's (1933, p. 193) literary and psychologically astute description, the outward appearance of the face is used as an index of the inward character.

Although our initial attitudes toward others in real-life interpersonal situations are surely affected—perhaps even generated—by clothing, speech, mannerisms, and facial expression, laboratory studies show that facial features alone can elicit a variety of subjective responses (e.g., Goldin, 1979; Secord, 1958; Secord, Bevan, & Dukes, 1953; Shoemaker, South, & Lowe, 1973). For example, the readiness with which subjects judge a pictured person's "intelligence" or "honesty" when the only stimulus available is a photograph of a face suggests that this is a familiar mode of social cognition for them. In addition, facial features elicit similar responses from subjects; that is, a particular face often evokes consensual responses from observers. Thus, it is consistent with both common experience and laboratory outcomes to assume that facial features may be believed to function as a visible index of behavioral dispositions and that many observers within a culture will agree about the "meaning" of particular facial configurations. Not only does Cassius have a lean and hungry look, but he also appears very bright and a bit dishonest, and these particular dispositions are evident to many observers from his face alone.

Under normal circumstances, making inferences about enduring personality characteristics, behavioral intentions, temperament, and other aspects of a person's "character," on the basis of outward appearances, such as facial features is at best harmless and at worst can lead to embarrassment and surprise. However, in the context of the criminal justice system, stereotyping on the basis of facial features could lead witnesses to select the wrong person in a police lineup and could also affect jurors' judgments of guilt and innocence (Shoemaker et al., 1973). On the latter point, Lown (1977) noted, "While more conclusive proof of this theory is needed, the probability that stereotyping is a factor that affects juries remains and is something with which lawyers should be concerned" (p. 87). Lown specifically addressed the question of a defendant's constitutional right to a fair trial and how this right may be compromised if jurors stereotype the accused on the basis of physical characteristics. Related arguments have been offered by Bull (Bull & Clifford, 1979; Bull & Green, 1980) and by McArthur (1982).

The research described in this report tested the hypothesis that faces can serve as stimuli that trigger consensual stereotypical responses in observers. Although this hypothesis has been generally confirmed in several early studies (Goldin, 1979; Hollingworth, 1922; Hull, 1928, pp. 111-138; Litterer, 1933; Rice, 1926; Shoemaker

The authors' mailing address is: Department of Psychology, University of Missouri, Columbia, MO 65201.


Copyright 1985 Psychonomic Society, Inc.

et al., 1973), it has received little systematic attention.[1] The present study demonstrates the existence of several criminal and noncriminal stereotypes that we attribute to faces.

## METHOD

### Subjects

One hundred and one undergraduate students (41 men and 60 women) at the University of Missouri were tested individually or in groups of 2 to 6 subjects.

### Apparatus and Stimulus Materials

The facial stimuli consisted of 100 white, middle-aged men's portraits cut from a casting directory (1971-1972 *Players Guide*). All photographs were achromatic, and were cropped to remove all extraneous information. Photographs with the following characteristics were used in the study: full-face camera angle, neutral expression, absence of facial hair, scars, glasses or headgear. A panel of judges (also drawn from the University subject pool), who had been instructed to look for and remove photographs of familiar actors, screened the portraits before they were used. Portraits were arranged in five fixed sets of 20 (5 across and 4 down) faces. A unique letter-number identification label was placed next to each portrait. The cardboard sheets containing the arrays of portraits fit onto the stage of an opaque projector.

### Procedure

Each subject received a resonse sheet that listed, in the following order, three criminal "occupations" (mass murderer, armed robber, and rapist) and three highly regarded (according to the results of a 1978 Gallup poll) professional occupations (medical doctor, clergyman, and engineer). Spaces for entering portrait identification labels were provided next to each of the six occupational titles. The subjects were instructed to choose one mass murderer, one engineer, one rapist, etc., from each of the 20-portrait arrays. The subjects were assured that each array actually did contain one, and only one, authentic representative of each of the six occupations. Additional instructions directed the subjects to use any face only once, and to select one face for every one of the six categories from each array (i.e., "don't know" responses were not permitted). From two to four of the possible five arrays were viewed by each of the subjects. Viewing time for each array was adjusted to allow all subjects to finish making their choices comfortably.

## RESULTS

The rationale of this study rests on the assumption that if the word (e.g., "rapist") either failed to elicit "pictures in the heads" of our subjects or did elicit a variety of idiosyncratic pictures, then the faces chosen by the subjects to represent the rapist should be dispersed randomly among the faces in each array. On the other hand, if the word "rapist" elicits consensually similar "pictures" in many of our subjects, then their choices of faces would be expected to be nonrandom, such that a smaller number of faces would attract a relatively larger number of responses. These predictions depend on the presence, within an array, of a portrait similar to the "picture" in the subjects' heads. Because all subjects viewed fewer than the total of five arrays, the responses from 56 to 58 subjects were included in each statistical analysis. While viewing an array, each of the subjects selected six portraits, one to match each "occupation." In this way, some subjects made 24 (6×4 arrays) selections, and other subjects made 18 selections (6×3 arrays).

The subjects were constrained from selecting a portrait for more than one occupation. Thus, the total number of eligible faces within an array decreased as the subject progressed from the first occupation to the sixth occupation; that is, the subjects searched through 20 faces for a mass murderer, 19 faces for an armed robber, 18 faces for a rapist, etc. Preliminary analysis of response frequency and position of portraits in the array indicated that the subjects were not matching faces and descriptors as a function of portrait position within an array.

The dependent measure in this study was the frequency with which subjects selected each face was an "occupational" exemplar. Thirty independent analyses were performed, one for each combination of array by occupational instruction condition (i.e., five arrays and six occupational instructions). Chi-square values were computed to compare the observed frequency of responding to each face in an array with the frequency expected if responses were randomly distributed across the faces. In 27 of the 30 analyses, selection of portraits was significantly ($p < .05$) nonrandom. Clearly, much consensual agreement exists among the subjects as to which faces are exemplars of the six occupations. Table 1 displays all the responses made to each face in Array 3 in all six instructional conditions. These data show that consensual responding was elicited when the subjects were asked to search for the various occupational exemplars. Thus, for example, when searching for the mass murderer among the 20 portraits, the subjects selected some faces much more frequently than would be expected by chance, whereas certain other faces were never selected. Table 1 also shows that in searching for the engineer in Array 3, the subjects demonstrated less consensual agreement in their responses than when they were searching for the mass murderer, a finding that was more or less repeated

**Table 1**
**Selection of Faces in Array 3 as Exemplars of Six "Occupations"**

| Portrait | Mass Murderer | Armed Robber | Rapist | Medical Doctor | Clergy-man | Engineer |
|---|---|---|---|---|---|---|
| 3A | 11 | 3 | 5 | 0 | 0 | 0 |
| 3B | 6 | 4 | 2 | 1 | 1 | 1 |
| 3C | 3 | 0 | 1 | 2 | 5 | 1 |
| 3D | 2 | 4 | 7 | 0 | 0 | 0 |
| 3E | 0 | 0 | 4 | 2 | 1 | 1 |
| 3F | 0 | 4 | 5 | 1 | 1 | 3 |
| 3G | 0 | 0 | 0 | 12 | 6 | 6 |
| 3H | 18 | 13 | 4 | 1 | 0 | 0 |
| 3I | 2 | 0 | 1 | 0 | 2 | 1 |
| 3J | 6 | 6 | 18 | 0 | 0 | 0 |
| 3K | 5 | 4 | 0 | 6 | 4 | 1 |
| 3L | 1 | 1 | 0 | 8 | 1 | 5 |
| 3M | 1 | 7 | 1 | 0 | 1 | 3 |
| 3N | 0 | 1 | 1 | 2 | 0 | 4 |
| 3O | 0 | 0 | 0 | 4 | 6 | 6 |
| 3P | 1 | 3 | 8 | 2 | 2 | 5 |
| 3Q | 1 | 0 | 0 | 5 | 20 | 6 |
| 3R | 0 | 5 | 1 | 2 | 1 | 2 |
| 3S | 0 | 2 | 0 | 4 | 3 | 4 |
| 3T | 1 | 1 | 0 | 6 | 4 | 9 |

in the other four arrays; the facial stereotype of the engineer appears to be less well formed than the stereotype of the other occupations included in this study. All three nonsignificant chi-square values were obtained from analyses of the subjects' responses to the "engineer" instruction.

The patterning of responses within the arrays suggested some of the reasons for the significant chi-square values. First, it was evident that some faces, when viewed in the context of other faces in the array, "attracted" large numbers of responses and other faces "attracted" only one or no responses. This finding suggests that some faces appear to be excellent exemplars of the various occupations, whereas other faces are excellent negative instances of a particular occupation. Unfortunately, because the 20 portraits within an array were always presented together, it is possible to evaluate only the relative and not the absolute potential of a face to evoke consensual responses.

Were subjects choosing faces to match their "pictures" of these occupations specifically, or were they choosing more affectively on the basis of how "good" or "bad" a face looked? To determine whether the subjects' choices of portraits were occupation specific, as opposed to simply "good-bad," we selected within each array one portrait with the highest total number of choices across all three "bad-guy" categories and one portrait with the highest total number of choices across all three "good-guy" categories. We reasoned that the face selected as the best exemplar of a bad guy would be chosen equally often as a rapist, robber, and murderer if the subjects were not making occupation-specific choices but rather choosing a face that looked "bad" (i.e., the null hypothesis). On the other hand, if choices were occupation specific, then distribution of responses to that portrait across the three categories would be unequal. Thus, for instance, the most often chosen bad guy should have been selected significantly more frequently as a rapist and not as a murderer or robber. Following this reasoning, separate chi-square analyses were computed on the data of each of the five most prototypical good guys and the five most prototypical bad guys. The distributions of responses to each of these 10 faces were tested against equal (33%) expected frequencies, which were based on the total number of responses made to each face. Chi-square values from the analyses of responses to the portraits of four of the five bad guys and two of the five good guys were significant [and one additional chi-square value for a good-guy face approached significance ($p < .10 > .05$)]. Extending this kind of analysis into the next five most frequently selected bad-guy portraits revealed that of all 10 prototype "bad" faces, responses deviated significantly from the expected equal frequencies in five cases and approached significance in one case ($p < .10 > .05$). Similarly, chi-square values for 3 of the 10 prototype good-guy faces were significant. These analyses suggest that, at least with some faces, the subjects were making discriminations among the three criminal types, and to a lesser degree they were also discriminating among the noncriminal professions.

## DISCUSSION

The present study clearly demonstrates that people are capable of processing perceptual stimuli by means of some kind of unintentional "matching-to-sample" technique, wherein the "sample" is a prototypic image or other memory representation of a face. Specifically, subjects seem to compare their perceptions of a face with a conceptual memory schema or template that represents their personal prototypic criminal or noncriminal face. In a sense, to use Lippmann's (1922) now familiar phrase, they can be said to have "pictures in their heads." Although the details of these "pictures"—stereotypes—must differ from person to person, the "pictures" must share common themes; otherwise, our subjects would not have so frequently been in agreement when they selected some faces and excluded others as exemplars of criminals and noncriminals. Moreover, our data imply that these "pictures" are refined; they are more than mere broad, global schemata. Not only are criminal and noncriminal facial prototypes represented, but various subtypes of criminals and noncriminals must also be represented in these memory schemata. However, compared with their global counterparts, these subordinate schemata appear less efficient as organizers of perceptual stimuli, because they yield lower judgment consensus (see Rubin & Kontis, 1983, for a discussion of current schema theory as it applies to memory for common objects).

The data of this study are important additions to the literature on stereotypes for four reasons. First, social stereotyping appears to be readily elicited by large numbers of faces; for example, reliable evidence for stereotyping was found in all five photo arrays used in this study. Second, facial stereotyping is not an isolated or rare phenomenon among our subjects; consensual responses to portraits were elicited from more than half the subjects who viewed the arrays. Third, observed consensual agreement about which face was that of a criminal and which face was that of an upstanding citizen was exceptionally high. Related to this outcome was the fact that if, for example, a face was frequently selected as a criminal, that face was never simultaneously selected by other subjects as a prototypic good guy. Stereotypes were mutually exclusive.

Finally, the present findings are noteworthy because the results are remarkably consistent with data reported by other investigators whose procedures and facial stimuli differed from ours. For example, when instructed to inspect a number of portraits of middle-aged men and identify those individuals who were most likely and least likely to have committed a particular crime, Shoemaker et al.'s (1973) subjects consensually chose some faces as criminals and some as unlikely to be criminals. Shoemaker et al. also reported that some faces were selected frequently as an exemplar of one kind of criminal, and simultaneously were infrequently selected as exemplars of other kinds of criminals. In another study (Bull & Green, 1980), subjects were presented with 10 faces to match with 11 different crimes. Although distribution of choice among the faces did not differ among some crimes (rape, arson, etc.), among other crimes (mugging, robbery with violence, etc.), one face was selected more frequently than others. Bull and Green concluded that expectations—socially shared stereotypes—do influence subjects' choices. Other investigators (Shepherd, Ellis, McMurran, & Davies, 1978) reported that judgments of the attractiveness of a particular face were dependent on whether subjects thought the face displayed was that of a "lifeboat captain" or a "murderer." Not only did the subjects downrate the attractiveness of the "murderer's" face, but when they reconstructed the face from memory using the Photofit technique, the face of the "murderer" was made more unattractive than the Photofit face of the "lifeboat captain."

Goldin (1979), in one of the most comprehensive series of studies of facial stereotypes, had subjects make Identikit constructions to get at the content of criminal face stereotypes (rapist, armed robber, embezzler, bookie, thief, political terrorist). She concluded that "people do possess notions about the appearance of the 'ideal' armed robber, rapist, etc., that these stereotypic beliefs can influence the processing

of 'criminal' stimuli" (Goldin, 1979, p. ii). Additional evidence demonstrating the influence of facial stereotypes on social cognition and perception can be found in a variety of sources (e.g., Klatzky, Martin, & Kane, 1982, Experiment 1; Secord et al., 1953).

Briefly, there is a considerable body of data accumulated over many years and by many investigators that strongly suggests the existence of consensual perceptual schemata for both criminal and noncriminal faces and also schemata for subtypes—e.g., rapist, clergyman—within those larger classes. These cognitive schemata appear to be capable of exerting control over many forms of social perceptual behavior.

What is the connection between the research described here and the criminal justice system's ability to identify and apprehend criminals using fair procedures? The legal process of resolving crimes is vulnerable to the influence of stereotyping at two critical points: during all the preliminary identification procedures that single out the probable criminal, and in the courtroom itself. If many people share well-developed, consensual stereotypes of criminal faces, stereotypes potentially capable of influencing cognitive social perception and personal judgment, then both the witnesses called upon to identify suspects and the jurors hearing the evidence in court may unwittingly use a defendant's physiognomy to make a "snap judgment" (Schneider, Hastorf, & Ellsworth, 1979), either derogatory or complimentary. Our findings suggest the definite possibility that the degree to which a particular individual's face invites facial stereotyping may influence the outcomes of any legal process in which they become involved.

One final point deserves comment. As we have seen, many faces were selected as either good- or bad-guy exemplars, and some faces were never selected. Was there anything about these faces, especially those frequently chosen as criminals or upstanding citizens, that set them apart from the other faces? In a preliminary attempt to explore this question, we had a fresh group of subjects, naive to the "reputations" of the people in the portraits, rate a number of the most popular exemplars from each of the three categories (good, bad, and neutral) on 24 scales of the Osgood Semantic Differential (Osgood, Suci, & Tannenbaum, 1957). Statistical analysis (separate one-way ANOVAs for each scale) indicated that on eight of the scales (dirty-clean, sane-insane, brash-cautious, excited-calm, cruel-kind, vulgar-refined, bad-good, and friendly-unfriendly), subjects' ratings differed significantly as a function of the class membership of the portrait. Although much further research is needed, it is interesting to note that in every significant comparison the ratings were consistent with expectation; bad guys were dirtier, crueler, insaner, etc., than good guys.

## REFERENCES


Ashmore, R. D., & DelBoca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior.* Hillsdale, NJ: Erlbaum.

Bull, R., & Clifford, B. R. (1979). Earwitness voice recognition accuracy. In G. Wells & E. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives.* New York: Cambridge University Press.

Bull, R. H. C., & Green, J. (1980). The relationship between physical appearance and criminality. *Medical Science Law*, **20**, 79-83.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition.* Reading, MA: Addison-Wesley.

Goldin, S. E. (1979). *Facial stereotypes as cognitive categories.* Unpublished doctoral dissertation, Carnegie-Mellon University, Pittsburgh, PA.

Hamilton, D. L. (1981). *Cognitive processes in stereotyping and intergroup behavior.* Hillsdale, NJ: Erlbaum.

Hollingworth, H. L. (1922). *Judging human character.* New York: Appleton.

Hull, C. L. (1928). *Aptitude testing.* New York: World Book.

Klatzky, R. L., Martin, G. L., & Kane, R. A. (1982). Semantic interpretation effects on memory for faces. *Memory & Cognition*, **10**, 195-206.

Lippmann, W. (1922). *Public opinion.* New York: Harcourt Brace.

Litterer, O. F. (1933). Stereotypes. *Journal of Social Psychology*, **4**, 59-69.

Lown, C. (1977). Legal approaches to juror stereotyping by physical characteristics. *Law and Human Behavior*, **1**, 87-100.

McArthur, L. Z. (1982). Judging a book by its cover: A cognitive analysis of the relationships between physical appearance and stereotyping. In A. H. Hastorf & A. M. Isen (Eds.), *Cognitive social psychology* (pp. 149-211). New York: Elsevier North-Holland.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Chicago: University of Illinois Press.

Rice, S. A. (1926). Stereotypes: A source of error in judging human character. *Journal of Personnel Research*, **5**, 267-276.

Rubin, D. C., & Kontis, T. C. (1983). A schema for common cents. *Memory & Cognition*, **11**, 335-341.

Schneider, D. J., Hastorf, A. H., & Ellsworth, P. C. (1979). *Person perception* (2nd ed.). Reading, MA: Addison-Wesley.

Secord, P. F. (1958). Facial features and inference processes in interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior.* Stanford, CA: Stanford University Press.

Secord, P. F., Bevan, W., & Dukes, W. F. (1953). Occupational and physiognomic stereotypes in the perception of photographs. *Journal of Social Psychology*, **37**, 261-270.

Shepherd, J. W., Ellis, H. D., McMurran, M., & Davies, G. M. (1978). Effect of character attribution on photofit construction of a face. *European Journal of Social Psychology*, **8**, 263-268.

Shoemaker, D. J., South, D. R., & Lowe, J. (1973). Facial stereotypes of deviants and judgments of guilt or innocence. *Social Forces*, **51**, 427-433.

Strachey, L. (1933). *Eminent victorians.* New York: Modern Library.


## NOTE

1. The works of Rice (1926) and Litterer (1933) are included as examples of research on facial stereotypes primarily because those investigators were first to explore the implications of the then new term, stereotype. However, experimental procedures employed by both investigators inextricably confounded the effects of facial features, bodily appearance, and clothing. Indeed, Litterer (1933), in discussing his results, draws particular attention to the clothing worn by the people in the stimulus photographs: "The coarse dress probably distinguished this woman from those women bedecked with modern clothes, furs and articles of display" (p. 61). For a history of the concept of stereotype, see Ashmore and DelBoca (1981).

(Manuscript received for publication July 2, 1984).