

Facing Faces: Studies on the Cognitive Aspects of Physiognomy

Ran Hassin and Yaacov Trope
New York University

Physiognomy, the art of reading personality traits from faces, dates back to ancient Greece, and is still very popular. The present studies examine several aspects and consequences of the process of reading traits from faces. Using faces with neutral expressions, it is demonstrated that personality information conveyed in faces changes the interpretation of verbal information. Moreover, it is shown that physiognomic information has a consistent effect on decisions, and creates overconfidence in judgments. It is argued, however, that the process of “reading from faces” is just one side of the coin, the other side of which is “reading into faces.” Consistent with the latter, information about personality changes the perception of facial features and, accordingly, the perceived similarity between faces. The implications of both processes and questions regarding their automaticity are discussed.

There are some people whose faces bear the stamp of such artless vulgarity and baseness of character, such an animal limitation of intelligence, that one wonders how they can appear in public with such a countenance, instead of wearing a mask. (Schopenhauer, 1942, p. 63)

[T]hus it is a settled point, that physiognomy of bodies is the daily guide of every man, whether he knows it, acknowledges it, or not. (Lavater, 1797, p. 28)

Take a minute and imagine the face of someone you love when he or she is happy, joyous, and glowing. Try to imagine each and every muscle in his or her face and to memorize its position. The expression, most probably, is very warm and kind. Now try to imagine the exact same expression superimposed on Richard Nixon's face. That is, try to imagine every muscle of Nixon's face in the exact posture of the equivalent muscle in the face you have just imagined. This image is probably not as pleasant as the one you engaged just a few seconds ago. But why? What is it that makes the expression look so different? Expressions of faces are, after all, patterns of movements of muscles in the face and their resultant state (see, e.g., Ekman, 1978; Ekman & Oster, 1979). Thus, the trick of imagination you just performed, at least to the extent that it was successful, should have created the same lovable expression.

This research was supported in part by a grant from the Israeli Ministry of Education and by Grant 1 R01 MH59030 from the National Institute of Mental Health.

Amos Tversky was involved in this project from the moment it was conceived through crucial stages of gestation. He died, however, before its birth. Out of respect for his devotion to formulating and perfecting prose, we are reluctant to put his name as an author, but note that his contributions are certainly commensurate with authorship. We also wish to thank Thomas Alley, Eran Chajut, Paul Ekman, Ofer Fine, Nehemia Friedland, Dave Kenny, Ido Levitan, and Eva Walther for their helpful comments on drafts of this article, and Mina Zemach for conducting the poll regarding lay beliefs in physiognomy.

Correspondence concerning this article should be addressed to Ran Hassin or Yaacov Trope, Department of Psychology, New York University, 6 Washington Place, New York, New York 10003. Electronic mail may be sent to hassin@psych.nyu.edu or to trope@psych.nyu.edu.

Chances are, however, that it did not. And one of the main reasons for that difference lies in the structure of the faces: Even if your efforts to imagine the muscles' postures were completely successful, something in the faces themselves created a different impression.

People have believed in and practiced physiognomy, the art of reading traits from faces, since the time of ancient Greece, when Aristotle is believed to have written a lengthy treatise on physiognomy (see, e.g., Brandt, 1980; Zebrowitz, 1997).¹ One of the highpoints of the history of physiognomy was in the second half of the 18th century and the beginning of the 19th century, when Johann Caspar Lavater, a Swiss theologian and mystic, published his four-volume *Essays on Physiognomy*. Lavater's success was enormous, and despite several attempts to undermine it (see, e.g., Hegel, 1931/1967, pp. 337–372), his physiognomy played a significant role in the intellectual and popular climates of these times (for a comprehensive discussion of Lavater's work, see Shookman, 1993). Later in the 19th century, phrenology briefly took physiognomy's place as the art of reading traits, this time from protuberances on the skull. At the beginning of the 20th century, physiognomy enjoyed renewed popularity, this time taking on a more “scientific” nature. For example, various vocational institutes, such as the Merton Institute, which provided services to such companies as AT&T, used physiognomy as one of their main tools in assessing candidates (Brandt, 1980; Zebrowitz, 1997).

The belief in physiognomy is not relegated to the past. In a survey of a representative sample of Israelis, we asked 535 respondents, “Is it possible, in your opinion, to know an individual's true personality traits from looking at his or her face?” Thirteen percent thought that it is possible to know all of an individual's personality traits from his or her face, 26% thought that it is possible to know many traits, 36% thought that it is possible to know few traits, whereas 25% thought that it is not possible to

¹ The term “physiognomy” is used in the literature in two ways. In the first, physiognomy denotes the facial features and their spatial organization. We will use it here in the second sense, that is, to denote the reading of traits from faces; “physiognomic information,” will refer to the trait information read (or inferred) from a face.

know any traits from looking at faces. Our survey indicates, then, that 75% of the Israeli population believes in physiognomy, whereas only 25% think that physiognomy is impossible (the sampling error of the survey is 4%).

Going beyond lay beliefs, there are good reasons to assume that the face, and physiognomy, play an important role in social cognition. First, the face is almost always seen whenever an interaction takes place (e.g., in "face-to-face" meetings). This means that the face, as a source of information, has a great advantage: It is available in almost every social situation. Second, until quite recently in the evolution of the human race, facial features, unlike facial (or behavioral) expressions, could not be willfully altered: What you see, in other words, is what Mother Nature shaped, not what the person wants you to see. Third, what you see is relatively stable. Unlike emotions that come and go, with their unique facial expressions (see, e.g., Ekman & Oster, 1979), the structure of the face is relatively stable: What you see now remains the same in 5 min, tomorrow, and next month. Fourth, a considerable amount of research suggests that there are areas in the human brain that specialize in face processing (see, e.g., Kandel, Schwartz, & Jessell, 1991; Young & Bruce, 1991). It seems only reasonable to assume that if the mind devotes special brain resources to processing faces it will try to extract as much information from the face as it can.

RELIABILITY AND VALIDITY OF PHYSIOGNOMIC INFERENCES

Considerable experimental evidence suggests that people can and do infer personality traits from faces. Initial studies in the first decades of the 20th century (see, e.g., Anderson, 1921; Cook, 1939; McCabe, 1928, cited in Secord, Bevan, & Dukes, 1953; Thornton, 1943) were followed by Secord's seminal and comprehensive work in the 1950s and 1960s (see, e.g., Secord, 1965; Secord et al., 1953; Secord, Dukes, & Bevan, 1954; Strich & Secord, 1956) and by more recent work that has focused mainly on babyfacedness (see, e.g., Albright et al., 1997; Berry, 1990, 1991; Zebrowitz, 1997; Zebrowitz-McArthur & Berry, 1987). Taken as a whole, this research shows that the process of inferring traits from faces is highly reliable. That is, different judges tend to infer similar traits from given faces. Some studies (e.g., Albright et al., 1997; McArthur & Apatow, 1983; Secord & Bevan, 1956; Zebrowitz-McArthur & Berry, 1987) even take this direction one step further, showing that this interjudge agreement is cross-cultural, thus suggesting that the cognitive work of reading traits from faces has some universal characteristics.

The picture that emerges regarding the validity of physiognomic judgments is more ambiguous. Most of the earlier studies that examined this question concluded that there is no significant correlation between facial features or physiognomic inferences and the traits individuals actually possess. Based on a thorough examination of physiognomy and graphology, Cohen (1973) concluded that "No significant or replicable relationships could be found between physiognomic or graphological characteristics and the mean evaluations of target persons by their comrades or their self-descriptions in a personality questionnaire" (p. 134). Fifteen years later, in a review of the literature on physiognomic inferences, Alley (1988) reached a similar conclusion: "Scientific re-

search, however, has generally found little or no validity in physiognomy" (p. 172).

However, more recent evidence suggests that face-based impressions may sometimes be valid. Most of the evidence supporting the "kernel of truth" hypothesis in face-based judgments has been obtained by Berry, Zebrowitz, and their associates (e.g., Berry, 1990, 1991; Zebrowitz, 1997; Zebrowitz, Voinescu, & Collins, 1996). Thus, for example, Berry (1990, Study 1) asked students to report their impressions of their classmates (after 1, 5, and 9 weeks of the semester have elapsed), and used these impressions as the criterion with which she compared independent evaluations of the classmates' photographs. She found significant correlations between peer and photograph evaluations on three dimensions: power (partial $r = .37$), warmth (partial $r = .39$), and honesty (partial $r = .42$). Similar results emerged in Study 2, in which the criterion was students' self-ratings, and in a study by Bond, Berry, and Omar (1994), in which a weak positive correlation ($r = .20$) between the behavioral measures of deceptiveness and judgments of photographs was found.²

The disagreement between the earlier results, which suggest that physiognomic inferences may be invalid, and the later ones, which support the validity of such inferences, has not been resolved in the literature. One possible explanation for this discrepancy is that the later studies examined traits that are related to each other by means of a more general impression (such as being an attractive or a babyfaced adult), whereas the earlier studies examined relatively discrete traits that are not parts of a more global characteristic of faces. However, this and other possible explanations have not been empirically tested and, thus, we feel that the question of the validity of physiognomic inferences remains open and awaits a more conclusive answer.

THIN SLICES AND THINGS THAT ARE NOT SLICES

In a related field in social cognition, researchers have concluded that "thin slices of expressive behavior" produce consensual and mildly valid inferences (Albright, Kenny, & Malloy, 1988; Ambady & Rosenthal, 1992; Borkenau & Liebler, 1992, 1993; Levesque & Kenny, 1993; Watson, 1989; see also Depaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997). "Thin slices" may be short videos (with or without a vocal channel), situations in which raters and targets sit in the same room and look at each other, situations in which raters watch targets through a one-way mirror, and so forth. We would like to stress two major differences between the study of thin slices and that of physiognomy. First, whereas thin slices are slices of *behaviors*, the experimental study of physiognomy examines the effects of *static* facial signs. Second, whereas from an ecological point of view thin slices convey a minimal amount of information, from physiognomy's point of view they are very rich sources of data (for a similar argument, see Zebrowitz, 1997, pp. 46–47). Thus, for example, consider the

² Borkenau and Liebler (e.g., 1992, 1993) have examined inferences from still photographs and found similar results, that is, that inferences based on photographs are consensual and mildly valid. However, their research used photographs of full bodies, and they do not report any controls over the emotions conveyed in the faces. Thus, it seems that it is not physiognomy alone that is responsible for the inferences made by their participants.

"zero acquaintance" paradigm (e.g., Borkenau & Liebler, 1992), in which the raters view a 90-s video of the targets. In these videos the rater can see how each target enters a room and walks around it. Moreover, the rater can see different facial expressions of the target, how he or she looks (and sounds) when reading aloud, the way he or she dresses, and many more facts and behaviors that may tell a true (or false) story about the target and may create (or annul) consensus amongst the judges. None of this relatively rich behavioral information is given to the rater when he or she only sees a picture of the target's face, that is, when physiognomy is practiced.

The study of physiognomy and that of thin slices, then, are studies of two related, but distinct, domains: Whereas the first is the study of the informativeness and use of short expressive behaviors, the second is the study of the informativeness and use of static facial signs. Hence, conclusions drawn in the former field—and, in particular, conclusions about reliability, validity, and about later usage of the inferred information—are not necessarily true of the latter.

READING FROM AND INTO FACES

Unlike the science, or art, of physiognomy and unlike most of the relevant research in modern psychology, the present research does not try to examine the veracity of physiognomic inferences, nor does it try to examine and describe an exhaustive system that links characters to faces. Instead, the arguments presented above regarding the availability, stability, and uncontrollability of faces lead us to hypothesize that physiognomy is an integral part of social cognition, and to further hypothesize two general roles faces may play in it.

First, people use personality data read from the face as a source of information about a target. The personality information inferred from the face is accessible to the perceiver's cognitive processes and is used in the interpretation of other relevant information, as well as in later judgments, decisions, and evaluations. Moreover, because, according to our hypothesis, personality impressions and faces are strongly related, the process that leads to the extraction of physiognomic data and to the use of that information has some of the characteristics of an automatic process. Second, people use available personality data as a source of information about the face. Thus, data about "the person behind the face" may be used in the perception and interpretation of the face.

These two functions of faces are reciprocal: Faces may change personality impressions, and personality impressions may change the perception of faces. The first process may be termed "reading from faces" (RFF). In RFF, the face is used as a source of information from which the perceiver extracts data concerning personality. This information is used by the perceiver for the interpretation of other relevant information and in later processes such as judgments and decisions. We conducted a series of four studies to test the RFF hypothesis. Study 1 examines whether and how physiognomic information about a person affects the interpretation of verbal information about that person. Studies 2 and 3 examine participants' confidence in judgments based on physiognomy and compare it to the accuracy of their judgments. Last in this series of studies, Study 4 examines the influence of facial photographs on decisions and whether people can ignore physiognomic information when they are explicitly asked to do so.

The second framework hypothesis suggests a complementary process that might be termed "reading into faces" (RIF). In RIF, information about personality changes the perception of faces. Thus, it seems reasonable to speculate that none of the readers of this article thinks that Einstein's forehead is short. Is this because there is something about his forehead, or is it because of something we know about Einstein? We suggest that in this case what we know about Einstein shapes the way we perceive his facial features, at least to some extent. Come to think of it, without any knowledge of Einstein, he might be judged to be an amiable old man, who wanders around in his large castle in the country all day long, inventing things that never work, ruining everything he lays his hands on, and being a great nuisance to Aunt Jane, that practical and efficient old lady. How big is *this* man's forehead? Well, we claim it will be perceived as smaller than Einstein's.

The process of RIF is examined in a series of three studies. In Study 5 we examine whether people attribute different facial features to persons of different character. Study 6 examines the core hypothesis of RIF, that is, whether different verbal characterizations of the personality of a target change the perceived features of his or her face. One implication of the RIF hypothesis—that personality portrayals should alter the perceived similarity between faces—is examined in Study 7.

READING FROM FACES

Study 1: The Effects of Physiognomy on the Interpretation of Verbal Information

According to the RFF hypothesis, personality information read from the face may be used in the process of interpreting other available information hence affecting impressions, judgments, and decisions. The present study examines the effects of physiognomic information on the interpretation of verbal information and, consequently, the conditions that increase and decrease the use of physiognomic information.

The degree of ambiguity of information has been found to be an important determinant of how the mind incorporates and uses it (see, e.g., Dion, Berscheid, & Walster, 1972; Herr, Sherman, & Fazio, 1983; Kunda & Sherman-Williams, 1993; Schwarz & Bless, 1992; Stapel & Schwarz, 1998; Trope, 1986; Trope, Cohen, & Maoz, 1988; see a review of similar claims regarding stereotypes in Kunda & Thagard, 1996). For example, Stapel and Schwarz (1998) have shown that when a target is ambiguous, expectancy priming results in assimilative interpretation, whereas when the information about a target is mixed, expectancy priming fosters contrast. Trope (1986) has demonstrated that the influence of information about the situation on the interpretation of another person's behavior is greater when that behavior is ambiguous than when it is unambiguous. Comparable results have been obtained in research on attractiveness. For example, Dion and her colleagues (Dion et al., 1972) have demonstrated that ambiguous performances of attractive women are perceived as superior to the same performances of less-attractive women and that the use of attractiveness decreases when performance is unambiguous.

In the present study, participants were presented with both photographs and verbal information about target individuals. On the basis of the research cited above, we hypothesized that the use of physiognomic information would be greater when verbal infor-

mation is ambiguous, compared with when it is unambiguous. Study 1 was designed to test this hypothesis.

Method

Overview

The design was a 2 (ambiguity of verbal information: ambiguous vs. unambiguous; between subjects) \times 2 (photographs: powerful vs. weak; between subjects) \times 2 (two different scenarios; within subjects). Each page presented to participants had both a photograph and verbal information. Participants were told that they were going to "take part in a study that examines impression formation from verbal information" and were instructed to rate the target on several personality scales.

Participants

One hundred twenty (63 female, 57 male) 12th-grade Israeli high school students volunteered to participate in the study.

Materials

Photographs. Two pilot studies examined the personality traits conveyed by two sets of facial photographs. In the first pilot study, 32 participants rated 22 full-face photographs of American students (taken from Matsumoto & Ekman, 1988), and in the second pilot study, 29 participants rated 30 full-face photographs of Israeli citizens. Participants in both studies rated the photographs on 22 personality traits (see Appendix A) and on two additional scales measuring global characteristics of faces— attractiveness and babyfacedness—that were found in previous research to affect inferences made from faces (e.g., Dion et al., 1972; Berry, 1990).³ All photographs had a similar background and were taken from a similar distance, and target faces in both groups were judged to be neutral with respect to emotion. Replicating the results of previous studies, the reliabilities of participants' ratings of the stimulus persons' traits were relatively high in both pilot studies ($\alpha = 0.83$ and 0.84 , respectively; interjudge agreements = 0.13 and 0.15). A varimax rotation factor analysis of participants' ratings revealed three main factors. A Power factor, explaining 27% of the variance in the first pilot study and 31% in the second, was comprised mainly of dominance (factor loadings = $.80$ and $.81$, respectively), charisma ($.79$ and $.78$), and confidence ($.77$ and $.80$). A Kindness factor, explaining 18% of the variance in both pilot studies, was mainly comprised of pleasantness (factor loadings = $.81$ and $.72$), kindness ($.80$ and $.83$), and to a lesser degree honesty ($.69$ and $.78$). A Competence factor emerged as the weakest factor, explaining approximately 10% of the variance in both pilot studies. This factor was composed mainly of seriousness (factor loadings = $.80$ and $.85$), competence ($.62$ and $.65$), and ambitiousness ($.62$ and $.63$).

The mean ratings of the three prominent traits in the Power factor (confidence, charisma, dominance) were used to select two sets of photographs. The first set was of two women, one of whom was judged to be high on the Power factor ($M = 6.02$, $SD = 1.8$), whereas the other was conceived as being low on this factor ($M = 3.17$, $SD = 1.7$). These two photographs were coupled with the verbal set depicting a female target. The second was a set of two men's faces, one high in Power ($M = 6.05$, $SD = 1.7$), and the other low in Power ($M = 3.38$, $SD = 1.85$). This set was coupled with the verbal set portraying a male target.

Verbal information. Ambiguous and unambiguous behavioral information was selected on the basis of a pilot study. A group of 60 pilot participants (coming from the same population as that of the study itself) rated 11 ambiguous and 11 unambiguous behaviors on thirteen 9-point personality scales. The mean ratings of the three prominent traits in the Power factor (confidence, charisma, dominance) were used to select two sets of verbal information—ambiguous and unambiguous. The two unam-

biguous scenarios had high ratings on the Power factor and relatively low variance ($M = 6.60$, $SD = 1.85$, for one scenario; $M = 6.12$, $SD = 1.37$, for the other), whereas the ambiguous scenarios had moderate ratings of the Power factor and larger variance ($M = 5.40$, $SD = 2.38$, for one scenario; $M = 4.70$, $SD = 1.98$, for the other).

Thus, for example, the actor in the following text was judged to be ambiguous on traits related to the Power factor:

Ayelet has recently completed a B.A. in computer science and has begun working in a big firm as a member of a team that develops ground-breaking software. A few days ago there was a meeting of the team, where the status of the project in general, and Ayelet's work in particular, were discussed. While people who approved and disapproved Ayelet's work discussed her part of the project, she sat quietly, without saying a word.

Whereas adding just one sentence at the end of the previous text made the actor seem unambiguous on traits related to the Power factor:

Upon leaving the room Ayelet said to her friend that "these people are really stupid. I don't understand why do I have to listen to them."

Design and Procedure

The study was conducted in groups of 20–30 participants, and participants were randomly assigned to experimental conditions. Participants received a booklet, with instructions on the first page telling them that they were participating in a study investigating impression formation based on verbal information. They were also told that we were interested in their subjective impressions and that there were no right or wrong answers to the questions they would be asked. This page of instructions was followed by the first experimental page, which consisted of the photographs in the upper left corner and the relevant scenario in the center. On the following page, participants were asked to rate the target on thirteen 9-point personality scales, three of which related to the Power factor, and on an attractiveness scale. The next two pages were similar in structure, but presented the second scenario. Thus, each participant read and rated two scenario-photograph combinations. The order of presentation of scenarios was counterbalanced.

It is important to note that the only explicit mention of the photographs was in the texts themselves, as shown in the following example (taken from the scenario presented above):

Ayelet (in the above photograph) has recently completed a B.A. in . . .

Dependent measure. Participants were asked to rate the targets appearing in the verbal descriptions on thirteen 9-point personality scales, three of which related to the Power factor (self-confidence, dominance, and charisma). These three scores were averaged separately for each scenario, and the scores of the two scenarios were treated as a repeated measure. The other traits that participants rated were tender-rude, wise-stupid, sociable-loner, cold-warm, strong-weak, nice-hideous, serious-not serious, truthful-deceitful, kind hearted-mean, introvert-extrovert.

Results and Discussion

Participants' ratings were subjected to a mixed analysis of covariance (ANCOVA) with Ambiguity of verbal information

³ The selection of traits to be read from faces in the current studies should be based on traits that participants can and do infer from faces. To select such traits, 30 pilot participants (taken from the same population as that of the study itself) were asked to look at each photograph and write down the prominent personality traits characterizing it. The 22 most frequently mentioned traits were selected for the present study.

Table 1
Rating of Targets on Power Scales (Study 1)

Text	Weak photographs		Powerful photographs	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Ambiguous	5.28	1.87	6.12	2.21
Unambiguous	6.83	2.23	6.53	2.12

(ambiguous vs. unambiguous) and Photograph (powerful vs. weak) as between-subjects factors, and Scenario (the two different scenarios) as a repeated measures factor. The means and standard deviations of the Power factor are presented in Table 1. It is important to note that the pilot data showed the target photographs did not differ in babyfacedness, $t(31) = 0.90$, *ns*, and that the difference in attractiveness was marginally significant (confident photographs were judged as more attractive than unconfident photographs), $t(31) = 1.90$, $p = .07$. Hence, we used the attractiveness ratings that participants provided as a covariate in the following analyses. The attractiveness covariate was significant in the ANCOVA, $F(1, 115) = 10.34$, $p < .01$, but it did not interact with any of the factors or their interaction, all $F_s < 0.54$, $p_s > .50$. All the analyses presented below use the adjusted means, but the means we report are the raw means of participants' ratings.

The ANCOVA yielded the hypothesized interaction between the photographs and the verbal information factors, $F(1, 115) = 4.63$, $p < .05$. Thus, when ambiguous verbal information was accompanied by weak faces, it was judged as less powerful ($M = 5.28$, $SD = 1.87$) than when it was accompanied by powerful faces ($M = 6.12$, $SD = 2.21$), but the effect of faces was different in the unambiguous condition: When verbal information was accompanied by unconfident faces, it was judged as slightly more powerful ($M = 6.83$, $SD = 2.23$) than when it was accompanied by confident faces ($M = 6.53$, $SD = 2.12$). Further analysis confirmed that in the ambiguous text condition, the difference between the ratings of the scenarios accompanied by weak photographs and those accompanied by powerful photographs was significant, $F(1, 57) = 5.21$, $p < .05$, whereas the same effect in the unambiguous text condition was not significant, $F(1, 57) = 1.92$, *ns*. Unsurprisingly, the main effect of the Ambiguity factor was highly significant, $F(1, 115) = 20.13$, $p < .001$. The main effect of the Photograph factor was not significant, $F(1, 115) = 1.32$. The effects of both factors, and their interaction, on the traits not associated with the Power factor were not significant, $F_s(1, 115) < 1.22$, $p_s > .28$.

Thus, as hypothesized, the effect of physiognomic information on the interpretation of verbal information was greater when the latter was ambiguous, compared with when it was unambiguous. The traits that were read from the targets' faces were, in a sense, read into the verbal information, thereby changing its interpretation. Hence, for example, when Ayelet's face was confident and dominant, the fact that "she sat quietly, without saying a word" was interpreted as expressing power. However, when the exact same assertion was made when Ayelet's face was unconfident and nonconspicuous, it was interpreted as expressing weakness.

These effects cannot be explained by either attractiveness or babyfacedness. First, the photographs did not differ in their baby-

facedness, and attractiveness was used as a covariate. Second, if the effects were due to attractiveness or babyfacedness, they should have changed the interpretation of traits that were previously found to be related to these factors, such as wisdom (related to attractiveness; e.g., Dion et al., 1972) and warmth (related to babyfacedness; e.g., Berry, 1991). However, the photographs only altered the interpretation of traits associated with the Power factor, thus suggesting that neither attractiveness nor babyfacedness can explain our findings.

Study 2: Overconfidence in Physiognomic Judgments

Study 1 demonstrated that physiognomic information affects the interpretation of verbal information and, hence, judgments of traits. Another important characteristic of judgments, in which physiognomy might also play a role, is the confidence with which they are made. The literature on people's confidence in their judgments and decisions suggests that people are usually overconfident (e.g., Oskamp, 1965; Wagenaar & Keren, 1986), although, in certain situations, they are also underconfident (e.g., Lichtenstein, Fischhoff, & Phillips, 1982). Griffin and Tversky (1992) argued that the patterns of overconfidence and underconfidence observed in the literature may be explained by the hypothesis that people focus on the strength (extremity, salience) of the available evidence with insufficient regard for its weight (predictive validity, reliability). For example, in evaluating a candidate's abilities, we tend to focus on the strengths of the evaluations in his or her letters of recommendation without sufficiently weighting factors related to their predictive validity (e.g., how well the writer knows the candidate, how easily he or she praises other students, and so forth). If people are highly sensitive to variations in the strength of evidence and are not sufficiently sensitive to variations in its predictive validity, then judgments will be overconfident when strength is high and validity is low.

The current study, therefore, investigates the confidence that accompanies physiognomic judgments. Specifically, the arguments presented above regarding the importance of faces in social cognition suggest that faces are what Griffin and Tversky (1992) call strong stimuli. Given that even the strongest findings regarding facial validity suggest that faces are moderately valid, judgments involving physiognomy should create overconfident judgments.

Method

Overview

Participants viewed pairs of photographs. For each pair they were asked to choose which one of the two individuals appearing in the photographs works in a certain profession (e.g., psychologist) and then to rate how confident they were that they had made the right choice. For each pair, the profession about which participants were asked (target occupation) was the actual profession of one of the individuals appearing in the photographs.

Participants

Participants were 23 (13 female, 10 male) second-year psychology students from Tel Aviv University, who volunteered to participate in the study. The mean age of participants was 23.5 years.

Materials

Seven pairs of photographs of target people whose occupations were known to us were compiled. All photographs had a similar background and were taken from a similar distance, and target faces were judged to be neutral with respect to emotion. To overcome problems of base rates, we matched the photographs in each of the pairs on two dimensions—gender and age. The occupation of one target photograph in every pair was chosen as the target occupation, that is, the one about which participants were asked (see the *Design and Procedure* section below). The target occupations were psychologist (male photographs), graphic designer (male), optometrist (female), clerk (male), producer (female), electrician (male), and software programmer (female).

All photographs were subjected to a pilot study, in which 12 participants (8 women, 4 men) were asked to rate their attractiveness and babyfacedness, on 9-point bipolar scales ranging from 1 (*very attractive*) to 9 (*very repulsive*) and from 1 (*babyface*) to 9 (*mature face*), respectively. We compared the attractiveness and babyfacedness of two groups: the group whose occupations were chosen as target occupations and the group of the paired photographs. These two groups did not differ in attractiveness or in babyfacedness, all t s < 1.23, *ns*.

Design and Procedure

Each participant received a booklet that presented the seven pairs of photographs, each of which appeared on a separate page. The photographs were placed at either side of the upper part of the page. The question pertaining to the target occupation (e.g., "Which of the two men above is a psychologist?") appeared in the center of each page. The order of presentation of the pairs was randomized as was the side of the page (right or left) on which each target photograph appeared.

Participants were told that, after viewing each pair, they would have to choose which one of the target persons practices the target occupation. They were also told that there were correct and incorrect answers and a reward of 50 NIS (\$15) was promised to the participant who chose the most correct answers.

Participants were asked to indicate their level of confidence in each judgment on an 11-point likelihood scale ranging from 0.5 to 1.0. Prior to their completing the booklet, participants were told that 0.5 represented a mere guess, as it was the probability of being right with no knowledge, and that absolute confidence was represented by 1.0.

Results and Discussion

In line with our hypothesis, the judgments revealed a high degree of overconfidence: Participants' confidence was almost twice that of their accuracy. The mean confidence of judgments was .68, whereas the mean accuracy was only .37, $t(22) = 5.62$, $p < .001$. Almost all of our participants (20 of 23) manifested overconfidence in their judgments, and as evident in Figure 1, all of the photograph pairs produced overconfident judgments. Thus, the overconfidence created by photographs when people judge occupations seems to be very robust.

However, two surprising results emerged. The first pertains to participants' accuracy. Participants' success rate (37%) was significantly less than that expected by chance, $t(22) = 2.95$, $p < .01$. Second, as inspection of Figure 2 reveals, participants' accuracy and confidence were negatively correlated ($r = -.20$, $p < .05$). That is, the higher the confidence, the lower the accuracy. These two findings may suggest that the faces contained misleading cues on which participants based their answers. According to this reading, the more salient these cues were, the more participants relied on them for their judgments, and the more confident they became.

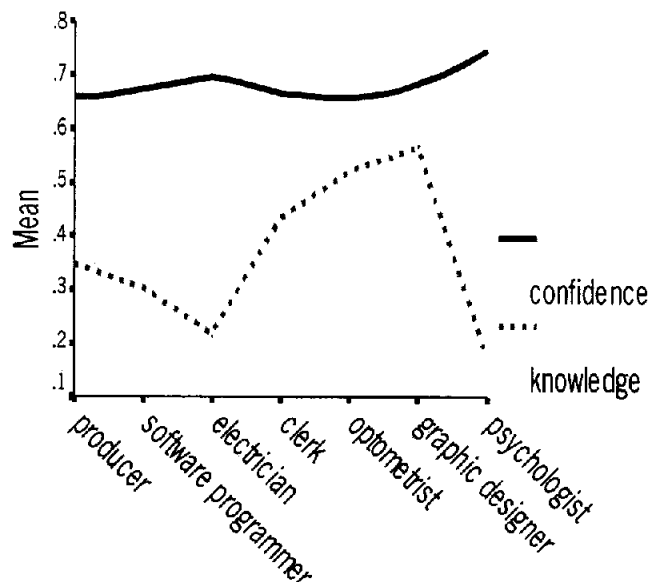


Figure 1. Confidence and accuracy of pairs.

However, because these cues were misleading, the more participants used them, the more erroneous they got, resulting in negative correlation between accuracy and confidence.

At least as far as the results regarding participants' accuracy is concerned, however, there may be an alternative explanation. Because we paired the photographs (so they would not differ on gender and age), one may argue that in pairing the photographs the experimenters inadvertently chose the photographs that would lead participants to err. Study 3 was conducted to examine this possibility.

Study 3: Overconfidence in Physiognomic Judgments: A Replication

Method

Participants

Twenty-three (10 female, 13 male) first-year students from the preacademic school of the Hebrew University volunteered to participate. The mean age of the participants was 22 years.

Materials and Procedure

The photographs in the current study were randomly paired. The random pairing was done separately for male and female photographs and resulted in seven new pairs of photographs. Each pair contained a photograph of one individual whose occupation served as the target occupation and a paired photograph. The group containing the photographs whose occupations were chosen as target occupations and the group containing the coupled photographs did not differ in attractiveness or babyfacedness, all t s < 1.23, *ns*.

Results and Discussion

The results replicated those of Study 2. Participants' mean confidence was .79, whereas their mean accuracy was .43, $t(22) = 8.08$, $p < .001$. Also, participants' success was signifi-

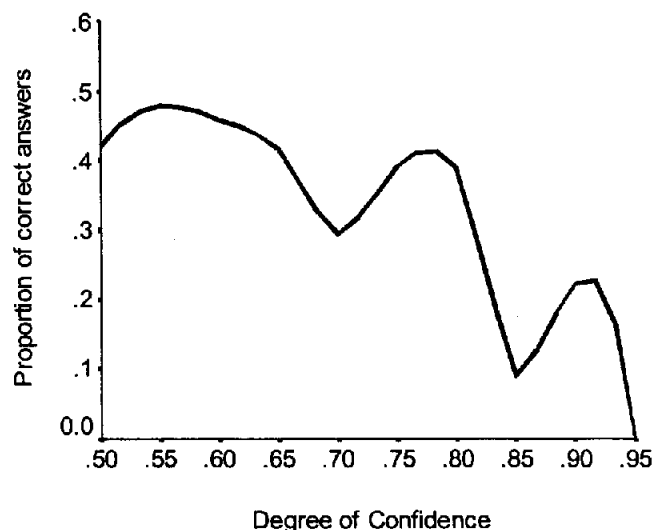


Figure 2. Accuracy as function of confidence.

cantly worse than chance, $t(22) = 1.86, p < .05$. As in Study 2, all of the pairs produced overconfidence, and all but one of our participants were overconfident. The correlation between accuracy and confidence was slightly negative but only marginally significant ($r = -.12, p = .07$).

In sum, two studies demonstrated that a high level of overconfidence accompanied judgments that relied on physiognomy. Although we think that the studies are not sufficient to conclude that faces in general are misleading, not even in our specific task, it seems reasonable to view them as supporting the claim that physiognomic inferences may not be valid.

Study 4: Physiognomic Effects on Decisions

In Study 1 we examined the process of integrating physiognomic and verbal information and uncovered some of its constraints. Studies 2 and 3 examined the confidence created by physiognomic judgments and demonstrated that faces create highly overconfident judgments. The current study examines the effect of physiognomic information on decisions. The hypothesis is simple: If physiognomic influence is strong enough, then it should also be detected at the level of decisions.

Study 4 was designed to test this hypothesis in one context in which physiognomy may have important consequences, namely, personnel selection. In a typical personnel selection situation, an interviewer is required to choose from different candidates for a certain job, basing his or her decision on several sources of information made available by a vocational institute (e.g., a battery of psychological tests, an interview, questionnaires that the candidates had completed). The interviewer must integrate these different sources of information and make a decision. Serious and thorough as the different sources of information about the candidates may be, they are still very limited and thus leave room for physiognomy to come into play.

The current experiment, using a laboratory simulation of personnel selection decisions, investigates two aspects of the relationship between physiognomic information and deciding which can-

didate will be hired for a job. First, it examines whether the personality information conveyed in faces influences such decisions. Second, it examines whether people can ignore physiognomic information when they are asked to do so.

Method

Overview

Participants simulated personnel-selection decisions. For each "job opening" they were presented with a pair of candidates and were asked to choose which one of the candidates will get the job. The presentation of each candidate included verbal evaluation and a facial photograph. The verbal evaluations depicted candidates that were relatively similar in terms of their suitability for the job. However, the faces of candidates were either representative or unrepresentative of the job. One half of the participants were explicitly asked to ignore the faces when making the decision, whereas the other half did not get any explicit instruction regarding the faces. The resulting design is a 2 (representativeness of faces) \times 2 (ignore instructions) between-subjects design.

Participants

Fifty-one (35 female, 16 male) Tel Aviv University students participated in the study. Thirty-one of the participants took part in the study for a course credit, whereas 20 were paid 15 NIS (\$4) each. Participants' ages ranged from 19 to 31 years, with a mean of 24.

Materials

Photographs. Recall that in each of the pairs of photographs in Study 2 one person was chosen more often than the other as the person who practices a certain profession. The person that was chosen more often, one might say, has a more representative face for that certain profession than does the other person. Out of the seven pairs of photographs in Study 2, we chose those in which the difference in the representativeness of the two photographs was the strongest. Thus, the psychologist pair led 83% of the participants to choose one of the photographs, whereas only 17% chose the other; 79% of the participants chose one of the clerk's photographs over the other, and 65% chose one of the two producer photographs. Let us call the photographs that were chosen by the majority of participants "representative photographs," and those that were chosen by the minority "unrepresentative photographs." Using the pilot data for Studies 2 and 3, we examined the attractiveness and babyfacedness of these two sets of photographs. There was no significant difference between the representative and unrepresentative photographs in babyfacedness, $t(11) = -1.43, ns$. The difference in attractiveness, however, was significant (representative photographs were more attractive than unrepresentative ones), $t(11) = 2.57, p < .05$.

Verbal information. For each profession, we composed several descriptions that were presented to participants as "short summaries of the Concluding Verbal Evaluations (CVE) of candidates, that were written by a vocational institute" (see an example in Appendix B). To avoid extreme differences in the suitability for the job conveyed by the CVEs, we conducted a pilot study in which 11 participants were asked to decide on the basis of the CVEs which of the candidates in each pair would get the job. Only pairs in which neither of the CVEs was chosen by more than 75% of the participants were included in the study.

Coupling photographs and verbal information. The experimental pages were A4-size pages, turned 90 degrees (so as to be wider horizontally). Each page was vertically divided into two equal halves, between which appeared a bold line. Each of the two CVEs of each pair appeared in the lower portion of one half of the page and above each CVE appeared a photograph. Thus, the resulting page included two parts, each containing

a CVE and a photograph. The coupling of photographs and CVEs of each pair was counterbalanced across participants, and the different pairs were presented in a randomized order.

Design and Procedure

Participants were randomly assigned to one of two groups. One group was instructed to "ignore the facial photographs of the candidates, so that your decision will reflect which candidate is more suitable for the job only according to the verbal description" (the ignore group). The other group did not receive any instructions concerning the use of the photographs (the no-instruction group). In both groups, CVEs appeared with either a representative or an unrepresentative photograph.

Participants took part in the study either individually or in small groups of 2–3 persons each. Each participant received a booklet, in which all three pairs of candidates appeared in a randomized order. Participants were given the following instructions:

In the booklet in front of you there are short descriptions and facial photographs of candidates, who applied for a job through a vocational institute. The job to which they applied is presented in the upper part of each page. The descriptions are short summaries of the Concluding Verbal Evaluations the candidates received upon finishing their tests at the vocational institute. Candidates appear in pairs, and you are required to decide, for each pair, which candidate you would have chosen for the job.

After making their decisions, participants were also asked to indicate, on a bipolar 9-point scale, the degree to which each of the candidates was suitable for the job. At the end of each booklet, participants were asked two more questions. First, they were asked to indicate ("yes" or "no") whether they thought the faces of the candidates had influenced their decisions. Participants who indicated that they were affected by the faces were then asked to assess to what degree, on a 9-point bipolar scale.

Scoring. Each participant was given a score ranging from 0 (*preferred the unrepresentative photograph in all three cases*) to 3 (*preferred the representative photograph in all three cases*).

Results and Discussion

If the type of photograph coupled with CVEs does not influence participants' decisions, then their mean scores should not deviate from chance. However, participants in both groups significantly preferred the candidates with representative faces ($M = 1.73$, $SD = 0.43$), $t(50) = 2.83$, $p < .01$. This tendency was significant both in the ignore group ($M = 1.75$, $SD = 0.18$), $t(23) = 2.3$, $p < .03$, and in the no-instruction group ($M = 1.71$, $SD = 0.22$), $t(26) = 1.74$, $p < .05$. The difference between the means of the two groups did not differ, $t(50) = 0.30$, *ns*. Thus, our two hypotheses were confirmed. First, when our participants simulated decisions regarding personnel selection, physiognomic information influenced decisions about candidates' careers. Second, participants were unable to ignore physiognomic information, even when asked to do so.

A 2 (instructions: ignore vs. no instructions to ignore; between subjects) \times 2 (faces: representative vs. not representative; within subjects) ANOVA performed on the suitability ratings of the different candidates revealed a similar finding (see Table 2). When CVEs were presented with representative faces, the candidates were perceived as more suitable for the job than when they were presented with unrepresentative faces, $F(1, 49) = 5.73$, $p < .05$. The effect of instructions, as well as the interaction between

Table 2
Ratings of Candidates on Suitability for Job Scale (Study 4)

Face	Ignore		No instructions	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Unrepresentative	6.18	1.54	6.35	1.54
Representative	6.71	1.52	6.60	1.49

Note. Ignore = group of participants instructed to ignore the facial photographs of candidates; no instructions = group of participants given no instructions concerning the use of the photographs.

instructions and representativeness, were not significant, all F s < 1.29 , *ns*.

Were participants aware of the fact that they used the physiognomic information? Seventy-two percent of the participants who were asked to ignore the faces said that they were successful in doing so, whereas 28% thought that the photographs influenced their decisions. There was no significant difference between the two groups regarding their decision pattern, $F(1, 49) = 1.11$, *ns*, which means that (a) awareness of the use of physiognomic information is not a necessary condition for this effect to occur and (b) awareness of this kind does not change the effect of physiognomic information on later decisions. In the no-instruction condition, only 40% of the participants thought that the photographs had no effect on them, whereas 60% thought they had. Here too, however, there was no significant difference between the decision patterns of the two groups. The participants who thought that the photographs influenced their decisions assessed their influence as small both in the ignore condition ($M = 3.57$, $SD = 1.91$) and in the no-instruction condition ($M = 3.80$, $SD = 1.85$). The two means do not significantly differ from one another, $F(1, 59) = 1.18$.

The effect of physiognomy on evaluations and simulated decisions resembles the effects of global characteristics of faces, such as attractiveness (e.g., Dipboye, Fromkin, & Wilback, 1975) and babyfacedness (e.g., Zebrowitz, Tenenbaum, & Goldstein, 1991). As revealed by our pilot data, the present findings cannot be explained in terms of babyfacedness because the group of representative faces did not differ in this respect from the group of the nonrepresentative faces. However, the results of the present study may be partially due to the slight but significant difference in the attractiveness of the representative and unrepresentative faces. Having admitted that, we would like to note, however, that it might be the case that the differences in attractiveness between the two groups are due to differences in physiognomic information, and not vice versa. People with positive physiognomic information, we argue, might be perceived as more attractive than people with less positive physiognomic information, in the same manner in which a photograph of a person is judged to be more attractive when the person is described as kind (Nisbett & Wilson, 1977).

In summary, the RFF hypothesis was examined in a series of four studies. The two pilots for Study 1 provide evidence for the reliability of physiognomic judgments in a hitherto unexamined culture: the Israeli. Study 1 demonstrates that physiognomic information changes the interpretation of other available information, namely verbal, and affects judgments especially when the latter is ambiguous. Studies 2 and 3 examine the confidence

created by physiognomy, showing that physiognomic information creates high degrees of overconfidence. Lastly, Study 4 demonstrates that physiognomic information may affect decisions simulating personnel selection. Moreover, this study shows that the use of physiognomic information has one important characteristic of an automatic processes, namely, perceivers cannot ignore physiognomic information, even when they are explicitly asked to do so.

READING INTO FACES

The flow of information, according to our hypotheses, is bidirectional: from faces to impressions (RFF) and from impressions to faces (RIF). The next series of studies examines RIF. Study 5 examines whether people attribute different facial features to individuals with different personalities. Study 6 examines the core of the RIF hypothesis, namely, whether information about others' personality changes the perception of their facial features. Finally, Study 7 examines the implication of RIF for perceived similarity of faces. Specifically, if information about personality changes the perception of facial features in a consistent way, then information about personalities should also change the perceived similarity between faces.

Study 5: The Facial Features People Associate With Personality Traits

To examine whether people attribute different facial features to individuals with different personalities, we composed four short texts depicting four kinds of individuals (kind hearted vs. mean, wise vs. stupid) and asked participants to read them and rate the facial features of the individuals described. Participants were not given any facial photographs, and thus any consistent differences between the facial features attributed to kind hearted versus mean or to wise versus stupid persons would serve as supporting evidence for the hypothesis that people attribute different facial features to individuals with different personalities.

Method

Participants

Nineteen (12 female, 7 male) second-year social science students at the Tel Aviv University volunteered to participate in the study. Their mean age was 24 years.

Materials and Procedure

Four short texts depicting four different individuals were used. Two of the texts were related to the Competence factor, describing one person as wise and the other as stupid. The other two texts were related to the Kindness factor, describing one person as kind hearted and the other as mean. The following are the short texts that portrayed the wise–stupid persons:

His colleagues claim that he is one of the wisest men they have ever encountered. He thinks very quickly, he is sharp, and he has outstanding analytic skills. Moreover, he is also very deep and thorough.

His analytic skills are extremely bad, and although it frequently seems that he does not understand what he is talking about—he likes saying what he has to say. "There is no doubt that he is one of the most stupid

men I have ever encountered," said one of his friends from high school.

Each participant received a booklet in which all four texts were presented in a randomized order. The participant was asked to read each text and then rate the person's face on twenty-nine 9-point scales, each pertaining to a different facial feature. Examples of such features are big eyes–small eyes, high forehead–short forehead, long ears–short ears, and so forth. There were two possible orders of the features: from the top of the face to the bottom or vice versa. Half of the participants received each order.⁴

Results and Discussion

Nineteen features were found to significantly differ between the ratings of the kind-hearted person and that of the mean one. The features, their values, and the corresponding *t* values are presented in Table 3. Five features were found to significantly differ between the ratings of the wise person and that of the stupid one. The features, their values, and the corresponding *t* values are presented in Table 4.

The results show, therefore, that people associate specific facial features with specific personality traits, hence confirming our hypothesis regarding the bidirectional nature of the relations between personality traits and facial features. It is important to note that 2 features were affected by the competence manipulation and not by the kindness one and that 16 were affected only by the latter. One of the implications of this dissociation is that no single halo effect can explain the results on both factors and, therefore, neither attractiveness nor babyfacedness can fully explain the current findings.

Study 6: Changing Perceived Features of Faces

According to the RIF hypothesis, information about one's character changes the perceived physical features of his or her face. Now that we know what facial features are associated with different personalities (Study 5), the crucial next step in the examination of RIF is to determine whether different personality descriptions change the perception of these features.

Previous research on the effects of social information on the perception of facial characteristics has documented effects of "reading into attractiveness" (e.g., Gross & Crofton, 1977) and "reading into babyfacedness" (e.g., Zebrowitz et al., 1991). Thus, for example, Nisbett and Wilson (1977) manipulated how likable a college instructor was and found that this manipulation changed people's perceptions of his or her attractiveness. The difference between these phenomena and the current hypothesis, however, is that we hypothesize an effect of personality information on the perception of physical features of faces, whereas the research cited above demonstrates the effect of personality information on global, psychological characteristics of faces (for research on the effects of social information on perception of physical attributes such as size, see Stapel & Koomen, 1997; Wilson, 1968).

⁴ Despite the fact that the research on face perception has underscored the significance of such features as, for example, the distance between the eyes divided by the distance between the eyes and the mouth (Parks, Coss, & Coss, 1985), we decided to examine what might be called "natural features."

Table 3
Facial Features Affected by the Kindness Factor (Study 5)

Feature	Kind-hearted person		Mean person		Kind-hearted person (relative to mean person)	<i>t</i>	<i>df</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Quantity of hair	7.16	1.68	4.05	2.25	Abounding	4.79****	18
Kind of hair	6.33	2.14	4.50	2.38	Curly	2.27*	17
Shape of eyebrows	6.00	2.17	3.44	2.18	Round	3.33***	17
Length of eyelashes	7.06	1.43	3.79	1.96	Long	4.60****	18
Eye size	7.17	1.42	4.45	2.91	Big	3.34*	17
Eye shape	7.06	1.35	3.64	1.79	Round	8.73****	18
Height of cheek bones	5.91	1.91	3.56	2.10	Low	2.68*	17
Length of ears	6.06	1.80	4.17	2.33	Short	2.52*	17
Length of nose	4.22	1.35	5.94	2.29	Short	2.85*	17
Shape of nostrils	6.11	1.23	4.71	2.40	Round	2.43*	17
Corners of mouth turning	6.78	1.86	3.22	2.07	Up	4.10***	17
Fullness of lips	6.17	2.22	3.39	2.20	Full	3.85****	17
Color of lips	6.34	1.91	3.56	1.95	Red	3.65****	17
Width of chin	5.84	1.25	3.67	2.57	Wide	3.13**	17
Shape of chin	5.94	1.21	1.67	2.61	Round	4.34****	17
Length of face	3.39	1.24	6.11	2.70	Wide	3.48**	17
Fullness of face	3.22	1.43	6.39	2.15	Full	4.37****	17
Smoothness of skin	6.76	1.22	4.11	2.29	Smooth	3.99**	16
Color of face	3.27	1.21	5.61	2.09	Colorful	3.84**	17

* $p < .05$. ** $p < .01$. *** $p < .005$. **** $p < .001$.

Because the results of Study 5 indicated that the effect of the Kindness factor was much stronger than that of the Competence factor, we focused our efforts on the former. We examined RIF with two different groups of photographs: One group contained photographs that were unambiguous in the Kindness factor, and the other contained photographs expressing moderate kindness.

Method

Participants

One hundred twenty (72 female, 48 male) 12th-grade pupils from a high school in Jerusalem volunteered to serve as participants. Their age ranged from 17 to 18 years.

Materials

Photographs. On the basis of the results of the pilots to Study 1, we chose two photographs that were judged as relatively extreme ($M = 6.32$, $SD = 1.59$, and $M = 3.82$, $SD = 1.62$) and two that were judged as

moderate ($M = 4.72$, $SD = 1.88$, and $M = 4.76$, $SD = 1.95$) on the Kindness factor.

Verbal information. The texts were selected on the basis of a pilot study in which four descriptions of kind-hearted individuals and four descriptions of mean individuals were rated by eight participants on thirteen 9-point trait scales. Three of the traits were related to the Kindness factor (i.e., pleasantness, kindness, and honesty; see pilots of Study 1), and their mean was used as the kindness score. Two portraits at each extreme of the kindness dimension were selected ($M = 1.63$, $SD = 1.20$, for the kind texts; $M = 7.30$, $SD = 1.60$, for the mean texts).

The following are examples of the texts presented to the participants (translated from Hebrew). First, a text depicting a mean individual, followed by one depicting a kind-hearted individual:

"He is one of the meanest persons I have ever met," testify many of Itay's acquaintances. His friends note that he is extremely cynical, and that his critical sense of humor offends many of his acquaintances. "Despite the fact that he acknowledges this, he does not try to soften his words so as not to be so insulting," they say. And one of them adds: "I think he just enjoys seeing people squirm."

Table 4
Facial Features Affected by the Competence Factor (Study 5)

Feature	Wise person		Stupid person		Wise person (relative to stupid person)	<i>t</i>	<i>df</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Height of forehead	6.35	2.06	3.94	2.59	High	2.38*	16
Shape of eyebrows	5.82	2.19	3.76	2.19	Round	2.20*	16
Depth of eyes	6.13	2.19	3.86	2.13	Plunged	2.66*	16
Short ears	5.94	1.78	4.24	1.99	Short	2.16*	16
Corners of mouth turning	6.18	1.33	4.35	2.06	Up	2.70*	16

* $p < .05$.

"He is one the nicest people I have ever met," testify many of Itay's acquaintances. His friends say that his kindness is exceptional, and that he cannot say 'no' to any of his friends or family's requests. Moreover, they say that Itay is extremely friendly and welcoming. "His pleasantness and kindness are both very special and very rare," says one of his friends.

Design and Procedure

The information was presented in booklets. On each page, a photograph appeared in the upper left corner, and the verbal information appeared in the center. Participants were told that they were taking part in a study on "verbal and visual information processing." Each participant was randomly assigned to one of the four conditions created by our 2 (photograph: extreme vs. moderate kindness) \times 2 (verbal information: mean vs. kind hearted) design and judged two target persons. Participants rated the target persons on two different sets of scales, each appearing on a different page. The first consisted of thirteen 9-point personality trait scales, three of which pertained to the Kindness factor. The second set consisted of nineteen 9-point scales, pertaining to facial features. On an additional scale, participants rated the attractiveness of the photograph. The reason for the inclusion of the trait scales was twofold: First, we assumed it would motivate participants to read the texts. Were it not for these scales, participants could have completed the feature scales without reading the texts, and the study would have been unintelligible to them. Second, the scales serve as a manipulation check.

As in Study 5, the facial features had two possible orders: from the top of the face toward the bottom, and vice versa. There were also three possible orders of the personality traits. Each participant was randomly assigned to a combination of the two. The order of presentation of the stimuli pages was counterbalanced.

Results and Discussion

An examination of the manipulation check (i.e., of the inferred personality traits) revealed that the means of the three kindness-related traits of the kind-hearted targets ($M = 2.35$, $SD = 1.17$) were significantly lower than those for the mean targets ($M = 6.72$, $SD = 1.36$), $F(1, 116) = 8.32$, $p < .01$.

Participants' ratings of the features were subjected to a 2 (photograph: extreme vs. moderate kindness) \times 2 (verbal information: mean vs. kind hearted) ANOVA, with photographs as a within-subjects factor. A main effect of verbal information was found for five facial features, as can be seen in Table 5. Interestingly, there was no interaction between physiognomic extremity and verbal information, demonstrating that the effect of RIF was not constrained by the strength of the physiognomic information.

How many of the 19 facial features should have changed significantly as a result of the verbal manipulation? If the changes were purely random, then it would have been reasonable to expect a change in one feature at most (5% of 19 features). Therefore, the feature changes obtained in the current study were four times as many as those expected by chance.

It is important to note that participants had the photographs of the targets in the booklets in front of their eyes. Hence, we claim that the effects of RIF here are perceptual. The fact that there was no interaction between the photograph and the verbal factors suggests that the RIF phenomenon is even stronger than we expected: It occurred not only when target photographs conveyed moderate traits, but also when they conveyed relatively extreme ones.

Study 7: Changing Perceived Similarity Between Faces

Extant models of similarity (see, e.g., Markman & Gentner, 1993; Ortony, 1979; Ritov, Gati, & Tversky, 1990; Tversky, 1977) agree that features play an important role in determining the similarity between two (or more) objects. Thus Tversky's (1977) influential contrast model suggests that the similarity of two objects (or terms, pictures, etc.) A and B is a weighted function of three arguments: the features that A and B share, the features that are exclusive to A, and the features that are exclusive to B. Ortony (1979) agreed with Tversky and even took featural models one step further by suggesting that they may explain understanding of nonliteral expressions. Markman and Gentner (1993) claimed that certain features are more important than others when similarity is computed, but agree with the principle, namely that similarity is, to a great extent, a function of features.

Given that similarity is a positive function of the number of common features of the objects in comparison, and given that information about personality changes the perceived facial features in a consistent way (Study 6), it follows that when two faces have similar personality descriptions, they will be perceived as more similar to each other than when their descriptions are different. The hypothesized effect, according to which similar personalities lead to the perception of similar faces, may be viewed as mirroring the findings of Secord and his colleagues, according to which similar faces lead to similar personality impressions (e.g., Secord et al., 1954).

Table 5
Facial Features Affected by Description of Target as Mean or Kind Hearted (Study 6)

Feature	Kind-hearted person		Mean person		Kind-hearted person (relative to mean person)	F	df
	M	SD	M	SD			
Length of ears	6.65	2.01	6.14	2.28	Short	4.09*	1, 116
Shape of chin	6.47	2.14	5.95	2.28	Round	3.99*	1, 116
Fullness of face	6.46	1.88	5.68	2.22	Full	8.86**	1, 116
Width of face	6.96	1.81	6.25	2.21	Wide	7.74**	1, 116
Attractiveness	4.70	2.07	4.10	2.05	Attractive	4.14*	1, 116

* $p < .05$. ** $p < .01$.

Method

Participants

Forty-seven (34 female, 13 male) first-year students from the preacademic School of the Hebrew University volunteered to participate. The mean age of the participants was 22 years.

Materials

Photographs. To ensure that the pairs of faces used in the study would be somewhat similar, a pilot study was conducted in which the similarity of 11 pairs of male photographs was assessed. Eighteen participants (12 women, 6 men), coming from the same population as that of the study itself, rated the similarity of each pair on a 9-point bipolar scale. The results indicated that, in all cases, participants judged the pairs to be more different than similar (all mean ratings were below 5). The four pairs with the highest similarity ratings (i.e., those that were nearest to the mid-point of the scale) were selected for the study ($M = 4.28$, $SD = 2.05$; $M = 4.22$, $SD = 2.29$; $M = 3.78$, $SD = 2.07$; and $M = 3.61$, $SD = 2.15$).

Verbal information. The trait manipulated in the present study was kindness, and the descriptions used were those that had been examined in the pilot of Study 6. There were four descriptions of mean people ($M = 7.75$, $SD = 1.49$; $M = 7.38$, $SD = 1.41$; $M = 6.50$, $SD = 2.30$; and $M = 7.25$, $SD = 1.75$) and four of kind-hearted ones ($M = 1.63$, $SD = .74$; $M = 1.70$, $SD = .52$; $M = 1.79$, $SD = .74$; and $M = 1.50$, $SD = .93$).

Coupling verbal information and photographs. To create the similarity condition, the four kind-hearted texts were randomly divided into two pairs, as were the mean ones, thus creating four pairs of texts with the two texts in each pair resembling one another on the Kindness factor. To create a difference condition, each of the four kind-hearted texts was randomly paired with a mean text, thus creating four pairs with the two texts in each pair depicting opposite individuals. In both conditions, each pair of texts was coupled with a different pair of photographs. Each of the photographs within each pair appeared 50% of the time with each of the texts.

Design and Procedure

The photographs-texts pairs were presented on A4-size pages, with the photographs placed above the corresponding text. Participants were randomly assigned to one of the two conditions. They were told that they were taking part in a study on impression formation based on scant information. They were asked to read each story and then write down several traits they expected the character depicted in the story to have. Then, on the following page, they were asked some general questions (e.g., "Do you know any of the persons in the pictures?" "If you do—how well?" and so forth). The last question on this page asked participants to rate, on a 9-point bipolar scale, the similarity of the two photographs. The same procedure was repeated for all four pairs.

Results and Discussion

As hypothesized, the similarity between the photographs in the similarity condition ($M = 3.67$, $SD = 0.23$) was greater than that in the difference condition ($M = 2.76$, $SD = .08$), $F(1, 45) = 4.22$, $p < .05$. Together, the results of Studies 6 and 7 suggest that, as hypothesized, information about character changes the perception of faces and that this change is strong enough to modify perceived similarity between faces.

When one takes the pilot study into consideration, however, it seems that the similarity condition and the pilot study ($M = 3.97$, $SD = .33$) resulted in roughly equal similarity ratings, $F(1, 40) = 1.13$, ns , whereas the ratings in the difference condition were much lower than those of the pilot study, $F(1, 39) = 6.36$,

$p < .05$. It seems, therefore, that the difference condition is primarily responsible for the effect. This conclusion is only tentative, however, because the three "conditions" were not part of the same experimental design.

GENERAL DISCUSSION

In examining RFF we have demonstrated that physiognomic information changes the interpretation of relevant information. The more ambiguous this information is, the more perceivers use the face (Study 1). Physiognomic information is also taken into consideration when perceivers make decisions about other people. Furthermore, even when asked to, perceivers are unable to ignore people's faces while simulating decisions regarding personnel selection, although they are quite sure that they are able to do so (Study 4). Finally, physiognomic information makes us highly overconfident about our judgments: Our confidence in physiognomy-based judgments far exceeds the actual accuracy of these judgments (Studies 2 and 3). At the risk of being somewhat repetitive, let us note that although some findings (e.g., the effect on interpretation of verbal information) parallel those obtained with global characteristics of faces such as attractiveness and babyfacedness, the current results cannot be explained by these characteristics, because they were either experimentally, or statistically, controlled. And although other findings (e.g., overconfidence) resemble those obtained with thin slices of expressive behavior, they were obtained here from static facial features.

The argument for the importance of faces in social cognition has received strong support from early work by Secord and his associates (e.g., Secord & Bevan, 1956; Secord, Dukes, & Bevan, 1954) and from the more recent research program by Zebrowitz-McArthur, Berry, and their colleagues (e.g., McArthur & Berry, 1987; Zebrowitz, 1997). Both projects strongly confirm the reliability of judgments based on faces and the fact that there is cross-cultural agreement in these judgments. Zebrowitz-McArthur and Berry have also conducted extensive investigation of the validity and real-life consequences of babyfacedness and have found support for the claim that under certain circumstances face-based judgments are valid (see, e.g., Berry, 1990, 1991; Berry & Landry, 1997; Zebrowitz, 1997; Zebrowitz & Collins, 1997; Zebrowitz, Voinescu, & Collins, 1996).

Reading traits from faces has fascinated writers and scholars for many years, but it represents only one side of the coin. The other side of the coin is that information about personality is read into the face. Thus, there are systematic differences between facial features attributed to wise people and those attributed to stupid ones, as well as between those of kind-hearted and mean individuals (Study 5). On the basis of these differences, verbal information about personality changes the perceived features of faces (Study 6) and, hence, the perceived similarity between faces (Study 7).

The present findings, it seems to us, might underestimate the role of physiognomy in social cognition. As described here, the processes of RFF and RIF are reciprocal: Personality information read from faces is read into verbal information, and verbal information (and other nonphysiognomic information) is read into faces. However, our studies do not determine when RIF or RFF will take place. One possibility suggested by our data is that information flows from the relatively unambiguous to the rela-

tively ambiguous source. Therefore, the ambiguous face of a candidate equipped with outstanding recommendations, dozens of publications, and funded research is likely to be perceived as smart. On the other hand, an ambiguous curriculum vitae of a candidate with an extremely clever face is likely to be perceived as stronger than that of a candidate whose face conveys the utmost stupidity.

A possibility we assume to more adequately describe real-life situations and which we find to be theoretically more interesting, is that both sources of information participate in a dialectic process. Extrapolating only slightly from our results, the performance of a person whose face looks wise is likely to be perceived as better than the same performance of a person whose face looks stupid. Thanks to physiognomy, therefore, we now know the wise-looking person to be wiser than the stupid-looking person. This knowledge, in turn, is used as a frame with which we (re)interpret their faces. The wise-looking person is now known to be wise and, hence, looks even wiser; the stupid-looking person looks more stupid. These two sources of information are thus used in a dialectic perception/interpretation, in which each serves as the frame for the perception/interpretation of the other, but at the same time uses the other as its frame of perception/interpretation.

The same phenomena that presumably result from dialectical serial processes may also result from a parallel distributed process (PDP) (for recent implementations of such networks in social cognition, see Kunda & Thagard, 1996; Smith & DeCoster, 1999). A network in which focal and contextual information are both represented, and in which (the process toward) the end state can affect both kinds of information, will by definition yield reciprocal processes. A network of this sort will yield dialectical processes if the flow of excitation between two nodes (e.g., A and B) is bidirectional and if it works the same way in both directions (either A and B excite one another or they inhibit one another). In this case, excitation of Node A will result in excitation of Node B, which will enhance the excitation of Node A, and so forth. Whether in serial or in PDP implementation, the implication of the above argument is the same: The effect of physiognomy might grow when taking part in a dialectical process.

However, the effects of physiognomy are not necessarily limited to on-line processes, and their amplification can be even stronger in processes that involve memory and recurrent judgments/decisions: As argued previously, ambiguous verbal information about a target is interpreted by physiognomic information. A judgment, based on this interpreted information and on the physiognomic information, is made with a high degree of confidence. This (by now) quite confident judgment will, in turn, be stored in memory as information about the person judged. The stored judgment, as well as the stored interpreted verbal information, are likely to be recalled the next time an interaction with the target takes place. It is also likely that in this interaction, the individual's face (and hence physiognomic information) will be available. Thus, in this interaction the physiognomic information is present in many different places: in the stored judgment, the stored verbal information, and the face itself.

Both of these hypothesized processes—the dialectical perception/interpretation and those involving memory and recurrent judgments or decisions—suggest to us that our studies underestimate the role of physiognomic information in social cognition. We believe that these processes are of special relevance and impor-

tance in this day and age, in which long written manifests function primarily as scrap paper for our children to draw on. Physiognomic information is an important part of the message we receive via MTV-like news broadcasts and newspapers. And the ambiguity our leaders master to "speak to all audiences" creates ideal conditions for physiognomy to flourish (see Study 1).

A Note on Automaticity

The last 2 decades have witnessed a proliferation of studies on the automaticity of cognitive processes in general and of social-cognitive processes in particular (see, e.g., Bargh, 1989, 1996; Greenwald & Banaji, 1995; Higgins, 1996; Jacoby, 1998; Kahneman & Treisman, 1984; Kunda & Thagard, 1996; Logan, 1989; Posner & Snyder, 1975; Shiffrin & Schneider, 1977; Wilson & Brekke, 1994). But, as is often the case, this flourishing of research on automaticity eventually led to deconstruction: Automaticity is no longer viewed as a unitary or monolithic concept (see, e.g., Bargh, 1989, 1996; Blair & Banaji, 1996; Gilbert & Hixon, 1991; Kahneman & Treisman, 1984; Macrae, Bodenhausen, Milne, Thorn & Castelli, 1997). Bargh (1996), for example, calls the dichotomy between automatic and nonautomatic processes a "false dichotomy," asserting that "automaticity is a continuum and not an absolute, all or nothing, state" (p. 177).

Despite the fact that most of the studies presented in this work were not specifically designed to examine the question of the automaticity of physiognomic inferences, some of their implications do support the claim that RFF and RIF are located on the automatic side of the automaticity continuum. Thus, participants in Study 4, who were specifically asked to ignore faces of candidates when deciding whether to hire them, used physiognomic information no less than did participants that were not asked to ignore faces, hence, suggesting that RFF is uncontrollable. Moreover, more than 50% of the participants in Study 4 attested that physiognomy had no effect on their decisions. To the extent that participants' explicit answers are to be believed, this result suggests that RFF can proceed without awareness. Participants' answers might also indicate that the process of RFF may be unintentional. If this is not so, then more than half of our participants consciously and willfully performed RFF, only to deny later that it had any effect on them. Although we cannot rule it out, we find it implausible.⁵

Similarly, it seems unreasonable to assume that the participants in Study 6 consciously performed RIF. Conscious RIF would have required our participants to make a very unlikely guess, namely, that the experimenters were really examining whether verbal information about a person's personality changes the perceived features of that person's face. We find this suggestion unlikely, and therefore believe that the results of Study 6 are consistent with the claim that RIF may also be unintentional and unaware.

In sum, although none of the above findings is beyond reasonable doubt, we feel that together they are sufficient to suggest that RIF and RFF are, at least to some extent and in certain situations,

⁵ It may be recalled, that even the 40% of participants who were not asked to ignore the faces, and hence who had no apparent motivation to deny their impact, attested that the faces had no influence on their decisions.

automatic. A more systematic resolution of this issue remains open for further investigation.

Are We Trapped in Our Faces?

We have established that the scope of physiognomic influence is large. This evidence, taken together with the considerations presented above regarding the presumed amplification of physiognomic influence, might create the impression that we are trapped in our faces. Are we?

Turning to related areas of research to answer the question, it has been suggested that conscious, controlled processes of thought may play a significant role in minimizing the long-term contaminations (Wilson & Brekke, 1994) created by stereotypes. But this answer is limited in nature: The activation and operation of processes of this kind are dependent on many factors, such as cognitive resources (Devine, 1989), motivation (Brewer, 1988), and explicit task instructions (e.g., Bodenhausen, 1988). Moreover, even when our participants met some of these requirements and engaged in a conscious decision, they could not ignore the physiognomic information, even when they were specifically asked to (Study 4).

The above arguments suggest that the answer to the question posed in the heading of this section is yes, to some extent we are trapped in our faces, because controlled processes are very demanding, and even when these demands are met, controlled processes are not always capable of saving us from our faces. But we are not doomed yet: The work presented here suggests one relatively automatic, uncontrolled way in which the influence of physiognomic information might be diminished: reading into faces. Thus, one of the consequences of RIF is that strong verbal information changes the perception of the face and, hence, the physiognomic information derived from it. So it is possible that, in the long run, verbal and nonverbal information might change the physiognomic information conveyed by one's face and, hence, "untrap" it.

We, personally, find the prospects of not being trapped in our faces quite encouraging.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, 55, 387-395.
- Albright, L., Malloy, T. E., Dong, Q., Kenny, D., Fang, X., Winquist, L., & Yu, D. (1997). Cross-cultural consensus in personality judgments. *Journal of Personality and Social Psychology*, 72, 558-569.
- Alley, T. R. (1988). Physiognomy and social perception. In T. R. Alley (Ed.), *Social and applied aspects of perceiving faces* (pp. 167-186). Hillsdale, NJ: Erlbaum.
- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, 111, 256-274.
- Anderson, L. D. (1921). Estimating intelligence by means of printed photographs. *Journal of Applied Psychology*, 5, 152-155.
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 52-74). New York: Guilford Press.
- Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 169-183). New York: Guilford Press.
- Berry, D. S. (1990). Taking people at face value: Evidence for the kernel of truth hypothesis. *Social Cognition*, 8, 343-361.
- Berry, D. S. (1991). Accuracy in social perception: Contributions of facial and vocal information. *Journal of Personality and Social Psychology*, 6, 298-307.
- Berry, D. S., & Landry, J. C. (1997). Facial maturity and daily social interaction. *Journal of Personality and Social Psychology*, 72, 570-580.
- Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, 70, 1142-1163.
- Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology*, 55, 726-737.
- Bond, C. F., Berry, D. S., & Omar, A. (1994). The kernel of truth in judgments of deceptiveness. *Basic and Applied Social Psychology*, 15, 523-534.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, 62, 645-657.
- Borkenau, P., & Liebler, A. (1993). Consensus and self-other agreement for trait inferences from minimal information. *Journal of Personality*, 61, 477-496.
- Brandt, A. (1980). Face reading: The persistence of physiognomy. *Psychologist Today*, 90, 90-96.
- Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (pp. 3-53). New York: Academic Press.
- Cohen, R. (1973). *Patterns of personality judgments*. New York: Academic Press.
- Cook, S. W. (1939). The judgment of intelligence from photographs. *Journal of Abnormal and Social Psychology*, 34, 384-389.
- Depaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review*, 1, 346-357.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5-18.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24, 285-290.
- Dipboye, R. L., Fromkin, H. L., & Wilback, K. (1975). Relative importance of applicant gender, attractiveness, and scholastic standing in evaluation of job applicant resumes. *Journal of Applied Psychology*, 60, 39-43.
- Ekman, P. (1978). Facial signs: Facts, fantasies and possibilities. In T. Sebeok (Ed.), *Sight, sound and sense* (pp. 125-156). Bloomington, IN: Indiana University Press.
- Ekman, P., & Oster, H. (1979). Facial expression of emotion. *Annual Review of Psychology*, 30, 527-554.
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, 60, 509-517.
- Greenwald, A. G., & Banaji, M. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.
- Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, 24, 411-435.
- Gross, A. E., & Crofton, C. (1977). What is good is beautiful. *Sociometry*, 40, 85-90.
- Hegel, G. W. F. (1967). *The phenomenology of mind* (J. B. Baillie, Trans.). New York: Harper Torchbooks. (Original work published 1931)
- Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, 19, 323-340.

- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.
- Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 3–26.
- Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman (Ed.), *Varieties of attention* (pp. 29–61). New York: Academic Press.
- Kandel, E. R., Schwartz, J. H., & Jessell, T. (1991). *Principles of neural science*. New York: Elsevier Science.
- Kunda, Z., & Sherman-Williams, B. (1993). Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin*, 19, 90–99.
- Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, 103, 284–308.
- Lavater, J. C. (1797). *Essays on physiognomy* (C. Moore, Trans.). London: H. D. Symonds. (Original work n.d.)
- Levesque, M. J., & Kenny, D. A. (1993). Accuracy of behavioral predictions at zero acquaintance: A social relations analysis. *Journal of Personality and Social Psychology*, 65, 1178–1187.
- Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgments under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge, England: Cambridge University Press.
- Logan, G. D. (1989). Automaticity and cognitive control. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 52–74). New York: Guilford Press.
- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology*, 33, 471–489.
- Markman, A. B., & Gentner, D. (1993). Splitting the difference: A structural alignment view of similarity. *Journal of Memory and Language*, 34, 517–535.
- Matsumoto, D., & Ekman, P. (1988). [Japanese and Caucasian facial expressions of emotion and neutral faces; slides of faces].
- McArthur, L. Z., & Apatow, K. (1983). Impressions of babyfaced adults. *Social Cognition*, 2, 315–342.
- McArthur, L. Z., & Berry, D. S. (1987). Cross-cultural agreement in perceptions of babyfaced adults. *Journal of Cross-Cultural Psychology*, 18, 165–192.
- Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35, 250–256.
- Ortony, A. (1979). Beyond literal similarity. *Psychological Review*, 86, 161–180.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *The Journal of Consulting Psychology*, 29, 261–265.
- Parks, T. E., Coss, G. R., & Coss, C. S. (1985). Thatcher and the Cheshire cat: Context and the processing of facial features. *Perception*, 14, 747–754.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.
- Ritov, I., Gati, I., & Tversky, A. (1990). Differential weighting of common and distinctive components. *Journal of Experimental Psychology: General*, 119, 30–41.
- Schopenhauer, A. (1942). *Complete essays of Schopenhauer* (B. Saunders trans.). New York: Willey Book Company.
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.
- Secord, P. (1965). Facial features and inference processes in interpersonal perception. In R. Taguiri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp. 300–315). Stanford, California: Stanford University Press.
- Secord, P. F., & Bevan, W. (1956). Personalities in faces: III. A cross cultural comparison of impressions of physiognomy and personality in faces. *Journal of Social Psychology*, 43, 283–288.
- 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.
- Secord, P. F., Dukes, W. F., & Bevan, W. (1954). Personalities in faces: I. An experiment in social perceiving. *Genetic Psychology Monographs*, 49, 231–279.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.
- Shookman, E. (Ed.). (1993). *The faces of physiognomy: Interdisciplinary approaches to Johann Caspar Lavater*. Columbia, SC: Camden House.
- Smith, E. R., & DeCoster, J. (1999). Associative and rule-based processing: A connectionist interpretation of dual process models. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology*. New York: Guilford Press.
- Stapel, D. A., & Koomen, W. (1997). Social categorization and perceptual judgment of size: When perception is social. *Journal of Personality and Social Psychology*, 73, 1177–1190.
- Stapel, D. A., & Schwarz, N. (1998). Similarities and differences between trait concept and expectancy priming: What matters is whether the target stimulus is ambiguous or mixed. *Journal of Experimental Social Psychology*, 34, 229–246.
- Strich, T. M., & Secord, P. F. (1956). Interaction effects in the perception of faces. *Journal of Personality*, 24, 270–284.
- Thornton, G. R. (1943). The effect upon judgments of personality traits of varying a single factor in a photograph. *Journal of Social Psychology*, 18, 127–148.
- Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, 93, 239–257.
- Trope, Y., Cohen, O., & Maoz, Y. (1988). The perceptual and inferential effects of situational inducements on dispositional attribution. *Journal of Personality and Social Psychology*, 55, 165–177.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–351.
- Wagenaar, W. A., & Keren, G. (1986). Does the expert know? The reliability of predictions and confidence ratings of experts. In E. Hollnagel, G. Maneini, & D. Woods (Eds.), *Intelligent decision support in process environments* (pp. 87–107). Berlin: Springer-Verlag.
- Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence for a surprising convergence with self-report. *Journal of Personality and Social Psychology*, 57, 120–128.
- Wilson, P. R. (1968). Perceptual distortion of height as a function of ascribed academic status. *Journal of Social Psychology*, 74, 97–102.
- Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116, 117–142.
- Young, A. W., & Bruce, V. (1991). Perceptual categories and the computation of "Grandmother." *European Journal of Cognitive Psychology*, 3, 5–49.
- Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Boulder, CO: Westview Press.
- Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and Social Psychology Review*, 1, 204–223.
- Zebrowitz, L. A., Tenenbaum, D. R., & Goldstein, L. H. (1991). The

- impact of job applicants' facial maturity, gender and academic achievement on hiring recommendations. *Journal of Applied Social Psychology*, 21, 525-548.
- Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "Crooked-faced": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, 22, 1258-1269.
- Zebrowitz-McArthur, L., & Berry, D. S. (1987). Cross-cultural agreement in perceptions of babyfaced adults. *Journal of Cross-Cultural Psychology*, 18, 165-192.

Appendix A

List of Traits for Study 1

- | | |
|--------------------------------|-----------------------------|
| 1. Wise-Stupid | 12. Sociable-Loner |
| 2. Dominant-Not conspicuous | 13. Conceited-Humble |
| 3. Naive-Sophisticated | 14. Kind hearted-Mean |
| 4. Calm-Agitated | 15. Nice-Hideous |
| 5. Clever-Silly | 16. Deceitful-Truthful |
| 6. Charismatic-Not charismatic | 17. Interesting-Boring |
| 7. Happy-Sad | 18. Introvert-Extrovert |
| 8. Cold-Warm | 19. Ambitious-Not ambitious |
| 9. Strong-Weak | 20. Confident-Unconfident |
| 10. Tender-Rude | 21. Serious-Not serious |
| 11. Obedient-Defiant | 22. Shrewd-Simple |

Appendix B

Two Examples for Concluding Verbal Evaluations (CVEs) Used in Study 4

The following texts were presented to participants as the CVEs of the candidates who applied for the psychologist job (the texts are translated from the Hebrew original):

1. Ido was born in the north of Israel in 1966. He was an excellent student in elementary school, and a very good one in high school. During the interview, he mentioned the fact that he played a central role in the social life of his school. The exams of our vocational institute reveal that Ido has very good analytic skills, but that in the area of general knowledge he scored a little below average. His conceptualization ability is very good, but he is a little lacking in the area of communicating his ideas in the visual medium. He has a B.A. in psychology from the Hebrew University, and a masters degree from Tel Aviv University. When asked about the event that influenced his life more than anything else in recent years, Ido said it was the birth of his son. For the past couple of years he has been working in the immigration department of the Jewish Agency. Ido told us, that as part of his last job he handled many difficult cases, and this experience made him want to take the offered job as a psychologist. Ido noted, that he very much liked the feeling that he was able to improve the psychological well-being of other people, and that he thinks the current job offers an excellent opportunity to do just that.

2. Nir was born in Jerusalem in 1965. He was a good student both in

elementary and high school. According to Nir, he was very popular in high school, both among the teachers and the students. The exams of our vocational institute show that Nir's analytic skills are a little better than average, and that he possesses good general knowledge. His verbal understanding suits that of his peers very well. When asked about the event that influenced his life more than anything else in recent years, Nir said it was the death of his mother. In the test that included drawing it seemed that Nir has very good technical skills, but that his conceptual ability is a little lacking. He has a B.A. in psychology from Haifa University, and a masters degree from Tel Aviv University. During the past few years he has been working in the youth department of Jerusalem Municipality. Nir testified that he liked the job very much, and that to the best of his understanding he has performed very well. "The interaction with the children and the ability to influence their lives," he said during the interview, "are the main reasons for my application for the offered job." An evaluation from Nir's boss could not be obtained as he is currently abroad.

Received March 25, 1999

Revision received November 1, 1999

Accepted November 2, 1999 ■