Journal of Personality and Social Psychology
1980, Vol. 39, No. 5, 861-873

Perceptions of the Impact of Negatively Valued Physical Characteristics on Social Interaction

Robert E. Kleck and Angelo Strenta
Dartmouth College

Individuals were led to believe that they were perceived as physically deviant in the eyes of an interactant. Following a brief discussion, they commented on those aspects of the interactant's behavior that appeared to be linked to the deviance. The experimental arrangements were such that the interactant did not, in fact, perceive them as deviant. Persons who thought that they possessed negatively valued physical characteristics found strong reactivity to the deviance in the behavior of their interactant, whereas those with a more neutrally valued characteristic did not. An expectancy/perceptual bias explanation was advanced to account for these results, though experimental demand could also be viewed as a plausible interpretation. Study 2 provided more definitive data on the demand aspects of the instructions used in the first study and reaffirmed that both the expectancy and the demand explanations were plausible. Study 3 used a new set of instructions explicitly devised to permit a test of the competing explanations. In addition to replicating the important findings of Study 1, the results of Study 3 in combination with those of Study 2 strongly undermine a demand interpretation of the original results. In a fourth study, persons who had observed the behavior of the interactants in Study 1 via videotape also perceived greater reactivity to an imputed negative form of deviance than to a neutral one. Data from this last study support the notion that the results of Studies 1 and 3 reflect the operation of an expectancy/perceptual bias mechanism and tend to rule out a self-fulfilling prophecy dynamic.

It is now well established that negatively valued deviant physical characteristics such as obesity (Maddox, Back, & Liederman, 1968), orthopedic disability (Kleck, 1969), and facial deformity (MacGregor, 1974) are important in determining some nontrivial social outcomes. Such physical characteristics affect, for example, the nature of the impressions we form of individuals having them (Kleinke, 1975), the causes we assign to these individuals' behavior (McArthur & Solomon, 1978), and whether we choose to affiliate with them or not (Snyder, Kleck, Strenta, & Mentzer, 1979).

Although a stigmatizing physical characteristic is important in determining some aspects of our responses to another individual, our own behavioral dispositions, contextual factors surrounding specific interactions, and non-physique-related characteristics of the deviant person also have an effect on our behavior. Wright (1960) has observed, however, that physically deviant individuals often simply take it for granted that all of the behavior emitted by the persons interacting with them is causally linked to their own deviance. Davis (1961) and Goffman (1963) have likewise commented on the tendency for the physically stigmatized to articulate their interactions exclusively in terms of their stigma.

Considered in attribution theory terms, the tendency that Wright and others are postulating appears to make some sense. Kelley

This research was supported in part by National Institute of Mental Health Grant MH 29446 to the first author. The authors would like to thank Nannette J. Hart and Alice E. Watson, who served as confederates in Study 1, and Katherine R. Rackow and Sarah A. Riddle, who performed this task in Study 3.

Requests for reprints should be sent to Robert E. Kleck, Department of Psychology, Dartmouth College, Hanover, New Hampshire 03755.

 0022-3514/80/3905-0861$00.75

(1971), for example, argues that in attempting to arrive at some understanding of human action, the naive perceiver uses what he labels the principle of covariation. To the greatest possible extent, the perceiver systematically varies situations, people, and stimuli and observes when an effect in question occurs. When it comes to understanding the impact of a physical characteristic such as a missing arm on the behavior of others, the physically deviant individual is, however, at a particular disadvantage. That is, he or she cannot vary the presence or absence of the physical characteristic to test for its effects on others. In any face-to-face encounter, the characteristic is potentially implicated as a cause of the behavior of others, simply because it occurs in contiguity with those behaviors. Physically deviant individuals can still apply the principle of covariation to persons, and to the extent that their interactants treat them differently, they could conclude that the characteristic is not implicated. Interestingly enough, laboratory studies have shown that at least in initial encounters, the variability in behavior across physically normal individuals is less when they are interacting with a physically deviant person than when they are interacting with one who is physically normal (Kleck, 1968, 1969; Kleck, Ono, & Hastorf, 1966).

Attribution theory aside, we have little systematic evidence that persons who are physically deviant perceive this deviance as causally implicated in the behavior of others. The purpose of the first study reported in the present article, therefore, is to examine the likelihood that persons who possess deviant physical characteristics will see links between those characteristics and the behavior of an interactant. This question is examined under conditions in which the characteristic can, in fact, have no direct impact on the other's behavior. Thus, at some level, the experiment constitutes a test of Wright's (1960) proposition that physically deviant individuals will perceive a relationship between how they are treated and their physical characteristics, even when such a relationship does not objectively exist. The specific hypothesis being tested is that if the physical characteristic at issue is negatively valued (epilepsy or a facial scar), it will more likely be implicated in an interactant's behavior than if it is not stigmatizing (an allergy). The initial study also sought to explore the specific aspects of an interactant's behavior that would be attended to by a physically deviant individual when attempting to assess the nature of the impact of the deviance on an interactant's behavior. It was tentatively hypothesized that a person with an obvious physical defect (facial scar) would tend to focus on the gaze behavior of an interactant, whereas someone with a nonobvious physical defect (epilepsy) would attend to nonverbal and verbal indicators of anxiety and tenseness.

Study 1

Method

Subjects. Subjects were 25 females enrolled in a coeducational college in the northeastern United States. Ten were unpaid volunteers from an introductory psychology course, and 15 were paid subjects recruited from campus dormitories. One subject had to be dropped, because she did not consent to have the scar material applied.

Confederates. Each subject interacted with one of two persons who were presumed to be subjects like themselves. These two college-age females interacted with subjects in each of the experimental conditions described below. They were blind concerning the purposes of the study, the experimental manipulations, and the dependent measures. They were told only that they would be discussing a particular topic with a number of students and that part of their task was to be behaviorally consistent across the various interactions.

Procedure. When subjects arrived at the laboratory, they were given a set of written instructions that randomly assigned them to one of three conditions: allergy, epilepsy, or facial scar. These instructions informed them that they would be involved in a discussion with another female student and that the experimenter's interest was in whether that person's behavior would be affected by the physical condition attributed to them.

All subjects then completed a biographical questionnaire, which they presumed would be exchanged with the person with whom they were to interact. For the allergy and epilepsy groups, the second page of the questionnaire asked subjects to list any significant aspects of their medical history. The experimenter asked them to write either "a mild allergy that is under drug control" or "a mild form of epilepsy that is under drug control," depending on their random assignment to experimental conditions. The confederates actually saw only the first page of biographical information. Subjects in the facial scar condition were asked if a cosmetic preparation used to simulate facial disfigurement in dramatic productions could be applied to their face. As this material dries, it gives the appearance of

a healed scar, the size and extent of which can be controlled by the application procedures. The scar was placed on the subject's right cheek between the ear and the corner of the mouth and was of a size to be clearly noticeable in face-to-face interaction. The experimenter gave the subject a small hand mirror to confirm that an authentic-looking scar had been placed on her face. As she put the mirror down, he informed her that he would have to put a moisturizer over the scar to keep it from cracking and peeling off. In the process of "moisturizing" the scar, the experimenter removed it without the subject's knowledge.

After cautioning subjects not to talk about their allergy, epilepsy, or facial scar unless the other person mentioned it, the experimenter then brought the confederate, who was in a waiting room down the hall, into the room where the interaction was to take place. Two chairs were placed near the center, and the confederate always took the same seat. The subject, who had come from an adjacent room, was asked to take the other chair, and both were told that additional instructions would be played to them over the speakers placed along one wall. The speaker cabinets also contained video cameras and microphones, which permitted the making of candid records of the interaction.

The experimenter then went to an adjoining control room where he turned on the videotape equipment and the tape recorder, which contained the remaining instructions. These instructions specified the discussion topic (strategies people use in making friends) and asked the two participants to begin their interaction. The confederates had been instructed to initiate all discussions with the same comment but to not dominate the interaction.

After approximately 6 minutes, the experimenter entered the room and asked both individuals to return to separate rooms and complete the questionnaires that had been placed there. The questionnaire asked the subject to rate the confederate's behavior on five dimensions: amount of eye contact, degree of tenseness, amount of talking, degree of perceived patronization, and amount of liking for the subject as a person. In addition, she was asked to estimate how attractive the confederate found her to be and whether the confederate would have been more comfortable with a closer or more distant seating arrangement. All ratings were done on 14-point bipolar scales.

When the subject had completed this form, she was told that she and the other individual had been videotaped and was asked to sign a release agreement allowing her part of the videotape to be viewed by others. The subject was then given the opportunity to view and comment on the videotape made of the other interactant's behavior. A split-screen recording format had been used, and the side of the video monitor on which the subject appeared was masked so that she could see only her interactant (the confederate). She was first shown a brief segment of the interaction to familiarize her with the format and then saw and heard (over earphones) the entire interaction sequence. She was instructed to comment as the tape was played, on any aspect of the other person's behavior that she felt was responsive to the manipulated physical state (i.e., scar, allergy, or epilepsy), and her comments were tape-recorded. The experimenter reiterated that the manipulation may or may not have had an effect and that the subject should not feel compelled to find effects where none were evident.

At the conclusion of this commentary, the subject was asked to summarize her comments in writing and make any additional observations. The general nature of the research was explained to her, and any questions that she had concerning the procedures were answered in detail.

Results

The first set of dependent measures was derived from the verbal and written descriptions that subjects gave of the confederates' behavior. The verbal commentaries were transcribed and appended to the written summary. Two judges, unfamiliar with the general purposes of the study or the specific physical characteristics manipulated, rated each description on the extent to which the person giving it felt that the physical state attributed to them had affected the confederate's behavior. These ratings were made on a 14-point bipolar scale, and the reliability across the two judges for the entire sample was .85. An analysis of variance based on the average judge rating assigned to each subject revealed a significant treatment effect, $F(2, 21) = 10.43$, $p < .001$. Individual comparisons showed that scar and epilepsy condition subjects were both judged to have made statements reflecting a greater impact of their characteristics on the confederates than were allergy subjects, $t(14) = 4.62$, $p < .001$, and $t(14) = 3.28$, $p < .01$, respectively. Scar and epilepsy were not different from each other.

A preliminary content analysis of the commentaries suggested that only comments concerned with gaze behavior (e.g., eye contact, looking, watching) and tenseness (e.g., uncomfortable, nervous, jittery) occurred with a frequency meriting statistical analysis. The experimenter and another person unfamiliar with the study independently coded the subjects' comments for the frequency with which these behaviors were seen as responsive to the manipulation of physical state. Interrater reliability coefficients were .98 for the gaze measure and .99 for tenseness. Analyses of variance of these frequency counts revealed significant overall treatment effects for tenseness, $F(2, 21) = 3.72$, $p < .05$, and for gaze,

Table 1
Mean Ratings of the Confederates on the Dependent Measures in Study 1

Dimension	Type of disability		
	Allergy	Epilepsy	Scar
Eye contact	12.00	10.37	11.75
Tenseness	2.37	7.12	8.00
Talking	9.37	9.37	10.37
Distance	7.37	7.87	8.62
Attraction	8.62	7.50	6.37
Liking	10.25	8.87	9.75
Patronization	1.87	3.25	4.12

Note. Ratings were made on 14-point scales; the higher the number, the higher the attributed level of the characteristic.

$F(2, 21) = 5.61$, $p < .02$. Between-groups comparisons on the tenseness measure demonstrated that epilepsy subjects used such terms more frequently in describing the confederates' behavior than did allergy subjects, $t(14) = 2.33$, $p < .05$, but no other comparisons were significant. The same comparisons for the gaze measure revealed a significant difference only between the scar and allergy groups, $t(14) = 3.3$, $p < .01$, with the former making more references to gaze behavior than the latter.

A multivariate analysis of variance, collapsed across confederates, was conducted for the seven dimensions on which subjects rated their interactant immediately following the encounter. This analysis yielded a significant main effect for experimental groups, $F(14, 30) = 2.95$, $p < .01$. Univariate analyses of variance conducted for each dependent measure demonstrated significant main effects for tenseness, $F(2, 21) = 15.59$, $p < .0001$; attraction, $F(2, 21) = 8.95$, $p < .002$; and liking, $F(2, 21) = 3.41$, $p = .05$. Specific between-groups comparisons for these three measures revealed the following pattern of results: Scar and epilepsy subjects perceived the confederate as significantly more tense than did allergy subjects, $t(14) = 6.08$, $p < .0001$, and $t(14) = 4.29$, $p < .001$, but scar and epilepsy groups did not differ from each other. Likewise, both scar and epilepsy subjects perceived that the confederate found them less attractive than did the allergy subjects, $t(14) = 3.79$, $p < .01$, and $t(14) = 2.26$, $p < .05$, though on this measure, the scar and epilepsy groups themselves differed, with the former rating themselves as less attractive to the confederate, $t(14) = 2.26$, $p < .05$. The significant main effect for the liking measure was contributed to primarily by the epilepsy subjects, who perceived that they were liked less well by the confederate than did the allergy subjects, $t(14) = 3.19$, $p < .01$. As the means in Table 1 show, the scar group perceived a level of liking by the confederates that was midway between the epilepsy and allergy subjects' perception, though not significantly different from either group.

The univariate analyses revealed no significant effects for the remaining four measures. An inspection of Table 1 does show, however, that the pattern of means for these other measures is generally consistent with the conclusion that scar and epilepsy subjects perceived their conditions as having a more negative impact on the confederates' behavior and disposition toward them as people than did allergy subjects.

Discussion

The general pattern of results is consistent with our hypothesis that a negatively valued physical characteristic is perceived by its possessor as having a greater impact on the behavior of an interactant than one that is not negatively connoted. Further, there is some evidence suggesting that the type of physically stigmatizing condition involved will affect the aspects of the other individual's behavior that are scrutinized for evidence that the physical deviance is being responded to. In the present case, for example, individuals who thought that they had a scar were more likely to focus on the gaze behavior of their interactants, whereas those who thought that the other attributed epilepsy to them were sensitive to behaviors indicating tenseness and anxiety.

These results are obviously consistent with an expectancy notion. Subjects presumably entered the experiment anticipating how others might respond to various forms of physical deviance and, when placed in interaction with a peer, readily found evidence consistent with these expectations. A questionnaire study (Kleck & Strenta, Note 1), using subjects

drawn from the same population as the experiment reported above, confirms that physically normal individuals do have clear expectations concerning the impact of various physical conditions on dyadic social interaction. Parallel to the findings reported here, individuals thought that a physically normal individual, when interacting with either a facially scarred person or a paraplegic in a wheelchair, would (a) be more patronizing, (b) find the individual less attractive, (c) be more tense during the interaction, and (d) tend to prefer relatively large interaction distances. When the physical condition was an allergy, perceivers expected it to have little if any impact on social interaction.

Although an expectancy explanation thus appears reasonable, we do not know whether the perceptions of the other individual's behavior were a direct function of what our subjects thought they would find or whether such thoughts served to modify the subjects' behavior, which in turn altered that of the confederates, with the modification being accurately reported by the perceiver. That interpersonal expectancies can have an impact on the behavior of others is now well documented (e.g., Snyder & Swan, 1978), and in the light of this robust phenomenon, we made a strong effort to control our confederates' behavior and kept them blind to the experimental hypotheses and conditions.

To further assess the possibility of confederate behavior differences as a function of subject expectancy, two independent female judges, age peers of our subjects, viewed the confederate's portion of the videotape for each of her interactions separately and rated her behavior on the dimensions that had been used as dependent measures for the subjects. Since interjudge reliabilities were quite low (.52 for the tenseness dimension being the highest), separate multivariate analyses of variance were conducted for each judge to determine if either could successfully discriminate the experimental conditions. Though the overall multivariate tests were far short of significance in both cases, univariate analyses were nevertheless conducted to see if either judge was successfully discriminating experimental conditions on any measure. Consistent with the multivariate findings, none of the individual tests approached significance.

As noted earlier, the scar and allergy manipulations of Study 1 were selected in part because questionnaire data (Kleck & Strenta, Note 1) suggested that persons expected the former to have important effects on interaction, whereas the latter was viewed as relatively trivial. Epilepsy was chosen as the third condition, because it, like facial deformity, is a negatively valued form of deviance (e.g., Kleck et al., 1968) but is nonobvious in much the same way as an allergy. Further, the allergy and epilepsy manipulations could be accomplished by the substitution of either term in otherwise equivalent instructions given to subjects. The scar manipulation on the other hand, involved the complex process of applying makeup to the subject's face, giving her an opportunity to view the scar in a mirror, and then removing it without her awareness. To some extent, therefore, the epilepsy condition provided a helpful control for the various conceptually irrelevant differences that of necessity existed between the scar and allergy conditions.

Given the direct parallels between the allergy and epilepsy manipulations, it appears reasonable to interpret the results as being a function of differential expectations that subjects had regarding the impact of these conditions on social interaction. The possibility exists, however, that the subjects themselves did not have clearly articulated differential expectations but attributed such expectations to the experimenter and were, therefore, simply acting consistently with perceived experimenter "demand." The fact that the study utilized a between-subjects design in which individuals were not aware of the other cells (physical conditions being manipulated) undermines such an interpretation, though not completely. In Study 2, we attempted to provide more direct data on experimenter demand by having individuals read the instructions that had been given to subjects in the allergy and epilepsy conditions in Study 1 and then respond to questions concerning their perceptions of the experimenter's hypothesis, as well as stating their own expectations regarding the likely outcome of the study. For comparison purposes, a third set of instructions was prepared, which was intended to more explicitly and un-

Table 2
Perceived Impact of Experimental Manipulations on the Behavior of Others in Study 2

Hypothesis	Experimental condition		
	Allergy[a]	Allergy demand[b]	Epilepsy[b]
Experimenter's	4.94	5.69	5.56
Own	3.17	3.38	5.06

Note. Ratings were made on a 7-point scale; high numbers denote perceived high impact.
[a] $n = 18$. [b] $n = 16$.

ambiguously communicate the experimenter's hypothesis.

Study 2

Method

Subjects and procedure. Fifty male and female college students from an introductory psychology course were randomly assigned to one of three experimental conditions. In each of these, subjects were given written instructions that they were told had been employed in a previous experiment that sought to explore the effects of a physical health condition on social interaction outcomes. These instructions placed them in the subject role of Study 1; they thought that the person they were to interact with had information regarding a health problem that they were experiencing (i.e., an allergy or epilepsy). The point of Study 1 as conveyed in these instructions was to have the subjects in that study assess the extent to which the imputed physical condition affected the other individual's behavior toward them.

When they had finished reading the instructions, subjects in Study 2 were asked to rate on a 7-point scale (a) the degree of impact that they thought the experimenter anticipated the imputed physical condition would have on the other individual's behavior and (b) the degree of impact that they personally felt such a physical state would have.

Experimental materials. The same instructions that had been used in Study 1 for the allergy and epilepsy manipulations constituted two cells of the present design. In both sets of instructions, the sentence most likely to affect perception of the experimenter's hypothesis was, "The question as to whether or not the other subject will behave differently toward you because he/she thinks you have an allergy (epilepsy) is an interesting but unstudied one." To provide a condition that would be less ambiguous regarding the experimenter's hypothesis and against which responses to these two sets of instructions could be compared, the above sentence was modified to read, "The question as to whether or not the other subject will behave differently toward you because he/she thinks you have an allergy is an interesting one and recent data we have collected suggest that the other subject's behavior will be different." This will subsequently be referred to as the allergy demand condition.

Results and Discussion

Table 2 shows the means for the subjects' ratings of what they thought the experimenter expected regarding the impact of the independent variable (physical condition) and what they personally expected the study to find. Consistent with our findings in Study 1, persons reading the epilepsy instructions reported that they expected a greater impact on another's behavior than did subjects reading either version of the allergy instructions, $t(32) = 3.8$, $p < .001$, for allergy versus epilepsy; $t(30) = 3.54$, $p = .001$, for allergy demand versus epilepsy. At the same time, however, persons in the allergy condition perceived that the experimenter expected this variable to have less of an impact than did persons in the epilepsy condition, $t(32) = 2.04$, $p < .05$. Subjects in the allergy demand group did not differ from epilepsy subjects in their attributions of experimenter expectation.

These data, rather than resolving the issue, leave open the possibility that the results of Study 1 are amenable to either an expectation or experimenter demand explanation, or perhaps some combination of the two. The data are helpful, however, in that they serendipitously provide the materials for a more definitive test of the two alternative explanations. The allergy demand instructions had been written to provide a condition for which subjects would have little difficulty inferring that the experimenter expected his manipulation to have an impact. The means in Table 2 reveal that this was indeed the outcome and that subjects perceived that the experimenter in the allergy demand condition expected as robust an effect as did the experimenter in the epilepsy condition. What is important, however, is that subjects in the allergy demand cell projected a hypothesis upon the experimenter (he was perceived as expecting high impact), which they themselves did not share (they expected relatively little impact), $t(15) = 5.57$, $p < .0001$. Thus, the allergy demand condition and the epilepsy condition were

equivalent in terms of perceived experimenter demand but very different in terms of subjects' expectations. If individuals' responses in the paradigm of Study 1 were a function of the inferred experimenter's hypothesis, then the allergy demand instructions should have generated results parallel to those found for the epilepsy condition of that study. If, on the other hand, subjects' own expectations were the important variable, the allergy demand instructions should have resulted in data parallel to the original allergy condition in Study 1. In Study 3, the allergy and epilepsy manipulations of Study 1 were replicated precisely, and the allergy demand instructions were included as a third condition.

Study 3

Method

Subjects. Subjects were 30 females from the same population as those used for Study 1. Fifteen were unpaid volunteers from an introductory psychology course, and 15 were paid subjects recruited from campus dormitories.

Confederates. Two college-age females served as confederates, and each was randomly paired with half of the subjects in each of the three cells of the design. The confederates were unacquainted with the subjects with whom they interacted and were blind to the purposes of the study, the experimental manipulations, and the dependent measures.

Procedure. When the subject arrived at the laboratory, she was given a set of written instructions that randomly assigned her to one of three conditions: allergy, allergy demand, or epilepsy. Throughout the experimental session, the experimenter was kept blind to the particular cell assignment of each subject, though he was familiar with the general nature of the research. The instructions informed the subject that she would be placed in a conversational encounter with another female student and that the experimenter's interest was in whether that person's behavior would be affected by a physical condition that was to be imputed to the subject.

Biographical information was exchanged as in Study 1, to lead the subject to believe that the other person perceived her as having an allergy or epilepsy. After the exchange, the experimenter brought the two persons together in the experimental room and told them that additional instructions would be played over a speaker in the corner of the room. This speaker concealed a video camera focused on the confederate and a microphone that was used to record the verbal exchange. The experimenter moved to the adjoining control room from which he played the tape-recorded instructions and videotaped the confederate. As before, the instructions specified "strategies people use in making friends" as the discussion topic.

Following approximately 6 minutes of interaction, both individuals were asked to go to separate rooms to complete a questionnaire. When the subject had completed this form, she was given the opportunity to view and comment on the videotape of the confederate's behavior. As in Study 1, she was asked to comment on any aspect of the other person's behavior that she felt was responsive to the medical information that had supposedly been given prior to the start of the interaction. At the conclusion of the commentary, which was tape-recorded, the subject was asked to summarize her observations in writing and make any additional comments. When she had finished writing, the experimenter consulted the subject's instruction sheet to determine which condition she had been assigned to. He then gave her the appropriate form containing the two questions that had been the focus of Study 2, that is, (a) what degree of impact she thought he anticipated the allergy (epilepsy) would have on the other person and (b) what her expectations had been regarding the degree of impact of the imputed condition. The rationale for the study was then explained to her, and any questions she had concerning the procedures were answered in detail.

Results

As in Study 1, the tape-recorded verbal commentaries were transcribed and appended to the written comments of the subject. Two judges, unfamiliar with the previous research and the general purposes of the present study, rated each description on the extent to which the person giving it felt that the physical condition attributed to her had affected the other person's (confederate's) behavior. As before, these judgments were made on 14-point bipolar scales, and the reliability between judges across the entire sample of 30 protocols reached an acceptable level, .73.

An overall analysis of variance based on the average judge rating assigned to each subject in each of the three conditions was highly significant, $F(2, 27) = 6.73$, $p < .005$. The mean ratings for each condition were 2.2, 2.45, and 5.0 for allergy, allergy demand, and epilepsy, respectively. As might be expected given this array of means, both allergy conditions were significantly different from the epilepsy cell, $t(18) = 3.37$, $p < .005$, and $t(18) = 2.52$, $p < .05$, for allergy versus epilepsy and allergy demand versus epilepsy, respectively.

A content analysis of the commentaries similar to that conducted for Study 1 revealed that references to tenseness and gaze behavior occurred with sufficient frequency across protocols to justify statistical analysis. Frequency

counts conducted by one judge and the experimenter for each of these categories were highly reliable: .96 for tenseness and .98 for gaze behavior. An overall analysis of variance of the frequency of tenseness references was significant, $F(2, 27) = 3.7$, $p < .05$, and separate comparisons revealed that as with the perceived impact data, the two allergy conditions differed from the epilepsy cell but not from each other, $t(18) = 2.61$, $p < .05$, for allergy demand versus epilepsy; $t = .0$, *ns*, for allergy versus allergy demand. An overall analysis of the frequency of gaze behavior references revealed that these did not discriminate the experimental conditions, just as they did not in Study 1.

As in Study 1, subjects completed a questionnaire that asked them to rate the confederate's behavior on seven dimensions: amount of eye contact, tenseness, patronization, amount of talking, liking, how attractive the confederate found them to be, and whether the confederate would have preferred a closer or more distant seating arrangement. All ratings were made on 14-point bipolar scales. A multivariate analysis of variance involving all seven measures, collapsed across confederates, did not yield a significant main effect for experimental groups, thus precluding univariate analyses of the dimensions separately.

Prior to debriefing, subjects were asked to rate the expectations which they had had and which they thought that the experimenter had had at the start of the session regarding the possible impact of the manipulation (allergy or epilepsy) on the other individual's behavior. This was done on the same 7-point scale used in Study 2. Unlike the subjects in Study 2, persons who had actually gone through the procedures did not project differential experimenter expectations as a function of conditions, $F(2, 27) = .48$, *ns*. This appears reasonable in that once one has experienced the relatively elaborate arrangements of the experiment, it is difficult to believe that the experimenter is expending all this effort for a manipulation that he believes will not have an effect. These responses could also be seen to further obviate an experimenter demand interpretation of the results, though since they were taken at the end of the experiment rather than prior to any interaction, they are somewhat problematic in this regard.

Subjects' own expectations regarding the impact of the manipulation were precisely parallel to those of subjects in Study 2. That is, allergy compared to allergy demand subjects did not anticipate a differential impact, $t(18) = 1.04$, *ns*, whereas both allergy groups differed from the epilepsy cell, $t(18) = 1.81$, $p < .10$, and $t(18) = 3.04$, $p < .01$, for allergy versus epilepsy and allergy demand versus epilepsy, respectively. As in Study 2, persons in the allergy conditions also perceived that the experimenter expected a greater impact for the manipulations than they did, $t(18) = 3.58$, $p < .01$, and $t(18) = 4.64$, $p = .001$, for allergy and allergy demand, respectively, whereas for the epilepsy condition, this was not the case, $t(18) = 1.54$, *ns*.

The results of Study 3 replicate those of Study 1 in that a negatively valued physical characteristic (epilepsy) is perceived by the person to whom it is imputed as having a greater impact on the behavior of an interactant than is one that is not negatively connoted (allergy). Further, these results, in conjunction with those of Study 2, strongly undermine the plausibility of a demand interpretation of the results and offer additional support for the expectancy notion advanced earlier.

As noted when the expectancy hypothesis was first advanced to account for the results of Study 1, there are two plausible mechanisms by which an expectancy might function in this paradigm. If a person to whom a negative characteristic has been imputed thinks it will have an impact on others, then he/she may modify his/her behavior in such a way as to directly affect the behavior of the other. From this perspective, the discrimination that our subjects were making between the allergy and epilepsy conditions could reflect real differences in confederate behavior created by a self-fulfilling prophecy dynamic. On the other hand, confederates may have been behaving in much the same manner across conditions, but subjects' expectations caused them to differentially perceive or interpret this behavior.

A preliminary assessment of these two possibilities was conducted as part of Study 1 by having naive judges view all of the videotapes

of the confederates in each of the three conditions in that study and rate their behavior on a number of dimensions. Though these judges were unable to discriminate the conditions, thus arguing against a self-fulfilling prophecy mechanism, it should be noted that the judges were uninformed regarding the experimental conditions or arrangements. It could be argued that one is less likely to detect true behavioral differences, particularly if they are relatively subtle, without knowledge of the situational constraints that have created the differences. Since our original subjects had direct knowledge of those constraints, they may have been better able to detect the impact of the variables on confederate behavior than were our naive judges. The fourth study has as part of its purpose an examination of this possibility, using correctly informed and misinformed observers as subjects.

The primary purpose of Study 4 was to assess whether observers of an interaction perceive the impact of a deviant physical condition on an interactant in a manner similar to actors who think that they possess that characteristic. It would be reasonable to expect, given the observers' more objective perspective on the interaction, that they would be less inclined to see an impact where none exists. At the same time, consistent with the results of both Study 1 and our questionnaire expectancy data, it was anticipated that observers would be less inclined to find behavioral responses to an allergy condition than to a facial scar.

Study 4

Method

Subjects. Subjects were 32 females drawn from the same population used in the previous studies. Twenty were volunteers from an introductory psychology course, and 12 were paid subjects recruited from dormitories.

Procedure. Subjects, who were tested individually, were told that their participation required that they be informed regarding a previous interaction between persons of their own age. They were then handed a description of this interaction, which randomly assigned them to either the allergy or scar condition. (The epilepsy manipulation was not used.) This description detailed the physical arrangements of the interaction, specified the true topic of conversation, and indicated how the experimental manipulation of physical condition had been achieved (i.e., biographical questionnaire in the case of allergy and cosmetic material in the case of facial scar). The subject was not told that the scar had been removed prior to the start of the interaction or that the allergy information had been deleted from the biographical sketch shown to the confederate.

Subjects were told that they would be looking at the videotape of one pair of subjects who had participated in the previous study. As in the first study, one half of the television monitor was masked so that the subject could see only the confederate. The subject was informed that this was being done to reduce the distraction that might otherwise be created and to allow her to concentrate on the "normal" interactant.

The videotapes in the present study were those made of the two confederates in the allergy and scar conditions of Study 1. All of the tapes from both of these Study 1 conditions were paired with a scar description for one half of the subjects and with an allergy description for the other half. Thus, one half of our observers had an expectancy consistent with those of the subjects in Study 1, and the other half had expectancies inconsistent with them.

When they had finished viewing the tape segment, subjects completed a set of self-report measures identical to those used in the first study. They also re-viewed the videotape and verbally pointed out specific aspects of the confederates' behavior, if any, that were responsive to the manipulation of the physical state of their interactant. As with the subjects in Study 1, they summarized these observations in writing. Finally, they were asked to respond to two questions: (a) "If you were interacting with someone who was facially scarred (had an allergy), how would you act so as to make the person think the scar (allergy) did not affect your behavior?" and (b) "In that interaction would you acknowledge or mention the scar (allergy)? Why or why not?"

As soon as the subject had completed her answers, the experimenter questioned her on her perceptions of the study and informed her of the deceptions involved. Any questions that she had regarding the procedures or the general research issues were answered in detail.

Results

The design of the study was a simple 2 × 2 analysis of variance in which observers were led to believe that they were viewing tapes of interactions involving someone with either a facial scar or an allergy (tape description), and this was crossed with whether the tape actually involved a subject who thought she had a facial scar or an allergy attributed to her by an interactant (type of tape shown). Main effects for tape description thus indicate that observer expectancies are important, whereas main effects for the type of tape shown suggest that there are actual behavior differences in the tapes that the observers were detecting independent of their expectations. The pres-

Table 3
Mean Ratings of the Confederates on the Dependent Measures in Study 4

Dimension	Scar description		Allergy description	
	Scar tape	Allergy tape	Scar tape	Allergy tape
Eye contact	7.25	8.62	8.00	6.62
Tenseness	7.62	7.87	6.87	5.50
Talking	10.62	11.75	10.00	10.50
Distance	8.37	9.00	7.25	8.75
Attraction	6.50	6.25	7.75	7.87
Liking	8.00	9.50	8.37	7.75
Patronization	6.12	5.75	4.87	5.12

Note. Ratings were made on 14-point scales; the higher the number, the higher the attributed level of the characteristic.

ence of these perceived behavioral differences would obviously support the notion that a self-fulfilling prophecy mechanism had been at work in Study 1 and that the scar and allergy condition subjects in that study were behaving differentially and thus eliciting different patterns of behavior from the confederates.

The verbal and written commentaries were used to derive the same measures as in Studies 1 and 3. Again, two judges unfamiliar with the design of the study independently rated each description on the extent to which it revealed (in the observers' eyes) some impact of the allergy or scar on the confederates' behavior. The interrater reliability was .79, and the average of the two judges' ratings was entered into an analysis of variance. Though judges tended to perceive the scar-described tapes as resulting in greater impact than allergy-described tapes, this tendency was only marginally significant, $F(1, 28) = 3.15$, $p < .10$. As in Studies 1 and 3, the observers' commentaries were content analyzed for all references to gaze behavior and tenseness by the experimenter and an independent judge, with respective interrater reliabilities of .97 and .99. The only significant effect for these two measures was that observers who thought that they were watching interactions involving a facially scarred person made more references to the confederates' gaze behavior as a probable outcome of the condition than did observers viewing persons with an allergy, $F(1, 28) = 30.94$, $p < .001$. There were no significant main effects or interactions involving type of tape shown to the observer.

The observers rated the confederate and her behavior on the same dimensions used by the subjects in Studies 1 and 3, that is, amount of perceived eye contact, tenseness, amount of talking, degree of perceived patronization, her apparent liking for her interactant, whether she would have preferred a closer or more distant seating arrangement, and how attractive she found her interactant to be. A multivariate analysis of variance of these ratings revealed that whether the interaction had been described to the observers as involving a scar or an allergy made a difference, $F(7, 22) = 3.94$, $p < .01$. In this analysis, there were no main effects for type of tape shown and no significant interaction between the description of the tape provided to the observer and the cell of the previous study (scar or allergy) from which the tape had been drawn.

Univariate analyses for each of the seven dependent measures revealed several effects of borderline significance. For example, subjects receiving the scar description rated the confederate as more tense than did those receiving the allergy description, $F(1, 28) = 2.88$, $p < .10$. As can be seen from the means for these measures, given in Table 3, the general pattern is for the scar to be perceived as having a more negative impact on an interactant's behavior and interpersonal disposition than an allergy has.

Finally, responses to the questions (a) "How would you act so as to make the person think the scar (allergy) did not affect your behavior?" and (b) "Would you acknowledge or mention the scar (allergy)?" were examined. Consistent with the other findings, 13 out of 16 subjects in the scar condition responded to Question a by suggesting that they would carefully monitor their gaze behavior, apparently because it was the most likely behavior to be perceived by the disfigured individual as linked to the scar. Comments such as, "I would try to look at her frequently, as if there were nothing wrong," "I would try to keep eye contact, not avoiding her face but at the same time not staring at it," and "I would attempt to meet her gaze and not avert my eyes from either it or the scar" are repre-

sentative of this concern. Only two persons in the allergy condition mentioned the need to monitor gaze behavior and, in general, unlike the scar condition subjects, persons in this condition disavowed the need for a conscious behavioral strategy. Comments such as, "It (the allergy) wouldn't affect my behavior, and I think I would feel perfectly relaxed around her and just be myself" or "I don't think that if a person did have an allergy I would act any differently toward her" are representative of this group.

Subjects in both conditions agreed (scar = 81%; allergy = 75%) that they would not explicitly mention the other person's physical failing unless that person brought it up. As might be expected, the reasons given for not initiating acknowledgment were quite different across the allergy and scar conditions. Of those persons in the allergy condition who stated that they would avoid mentioning the issue, six (50%) indicated that it was primarily because the condition was "uninteresting" or "irrelevant to the interaction." Only one individual felt that explicit mention of the allergy might generate embarrassment in the conversation. Of the 13 persons who said that they would avoid mention of a scar, on the other hand, 8 (62%) alluded to the embarrassment and anxiety that would probably follow from explicit acknowledgment.

Discussion

As expected, observers of a dyadic interaction perceived that if one member of the dyad was facially scarred, this would have a greater impact on the behavior of the other than if that person had an allergy. Though the general pattern of attributions made to the "normal" interactant by the observers in Study 4 was similar to that made by the actors in Study 1, the differences were clearly less robust. Perhaps the most striking similarity between the results of the two studies is that persons who think one member of a dyad has a facial scar (Study 4) or who think they themselves have that scar (Study 1) are very prone to see the gaze behavior of the normal interactant as reflective of responsiveness to the physical defect. It should be kept in mind that subjects in both studies were female and that there is some evidence that women are more responsive to the gaze behavior of an interactant than men are (e.g., Exline, 1963; Kleck & Nuessle, 1968). Further, gaze behavior appears to be an aspect of the behavior of others that we do not consciously attend to in most situations and, therefore, have little awareness as to what would constitute a normal level and style of looking (Ellsworth & Ludwig, 1972). If for some reason our attention is drawn to that behavior, we may conclude, quite erroneously, that it is linked to a specific aspect of our own person (as in Studies 1 and 3) or to a physical characteristic of one member of a social interaction (as in Study 4).

The results of Study 4 provide support for the notion that we are working with a perceptual bias rather than a self-fulfilling prophecy. That is, the different expectations that the actors brought to their interactions with the confederates in Study 1 did not appear to have a direct impact on the confederates' behavior. This conclusion is supported by the lack of a main effect for type of tape in Study 4. Observers in this study saw no more evidence for an impact of a facial scar on social interaction when they were seeing a tape of an interaction involving a person who thought that she had a scar than when they saw a tape of an interaction involving a person who thought that an allergy had been attributed to her. This is indirect evidence at best but is consistent with the data reported in Study 1 suggesting that judges could not discriminate differences in confederate behavior as a function of the nature of the expectancy that the confederate's interactant brought to the conversational interaction.

We have focused attention on this issue because it has been repeatedly demonstrated that how we expect others to behave can directly affect their behavior (Snyder & Swan, 1978). From this perspective, the causal chain would include these four links: (a) an expectancy or hypothesis for how an interactant will behave, (b) which leads to a change in own behavior, (c) whereupon the change in own behavior modifies the behavior of the interactant, and finally (d) the change in the interactant's behavior is perceived as confirming the expectancy rather than being linked to the change in own behavior. We had

created a paradigm that we hoped would rule out such a mediating mechanism, primarily by keeping our confederates blind to the experimental manipulations and by stressing the need for constancy in their behavior. It is obviously necessary to provide independent evidence for the success of this strategy, and this has been done in part through Study 4 and in the independent judge assessment of confederate behavior in Study 1.

One could question the importance of ruling out the self-fulfilling prophecy explanation for the present results. Whether the tendency to perceive a salient physical characteristic as affecting the behavior of those who interact with us is mediated by our own change in behavior, which in turn has an effect on those interactants, or is mediated by perceptual bias on our part, the result is essentially the same, that is, as physically deviant individuals we find evidence that our deviance is important to our social outcomes. Viewed from the perspective of the physically normal individual, however, which process is at work clearly could have quite different implications. If the interactive outcomes I experience with physically deviant individuals are largely a function of expectation-linked behaviors on their part, I can consciously not respond to those behaviors or strive to disabuse the deviant individual of his or her expectations and thus break the expectation–behavior interactive outcome cycle. If, on the other hand, it is simply a case of perceptual bias, then no matter what I choose to do, my behavior can be causally linked to the physical defects of the persons with whom I interact. Indeed, subjects in the fourth study who suggested that acting naturally was the best strategy to avoid having one's behavior linked to the physical failing of another nevertheless saw evidence of responses to a physically deviant characteristic in the "natural" responses of our confederates.

It should be kept in mind that the paradigm used in the first study is one in which persons are led to believe that they are temporarily deviant, and only in the eyes of one other person. The leap from these results, therefore, to the conclusion that they support Wright's (1960) assertion that physically handicapped persons are often prone to articulate their social reality entirely in terms of their handicap, even when the objective facts do not support such a construction, is tenuous at best. It is now necessary to demonstrate that persons who are permanently physically deviant make the same kinds of disability-linked attributions to a natural stream of behavior as did the subjects in the present studies. It is entirely plausible that with some history of interaction with others, a physically deviant individual develops strategies for testing and evaluating whether a specific behavior or social outcome can be correctly attributed to a physical failing. Research directed toward this goal is now underway.

Reference Note

1. Kleck, R. E., & Strenta, A. *Expectations regarding the impact of negatively valued physical characteristics on social interaction.* Unpublished manuscript, Dartmouth College, 1979.

References

Davis, F. Deviance disavowal: The management of strained interaction by the visibly handicapped. *Social Problems*, 1961, *9*, 120–132.

Ellsworth, P. C., & Ludwig, L. M. Visual behavior in social interaction. *Journal of Communication*, 1972, *22*, 375–403.

Exline, R. V. Explorations in the process of person perception: Visual interaction in relation to competition, sex, and the need for affiliation. *Journal of Personality*, 1963, *31*, 1–20.

Goffman, E. *Stigma: Notes on the management of spoiled identity.* Englewood Cliffs, N.J.: Prentice-Hall, 1963.

Kelley, H. H. *Attribution in social interaction.* Morristown, N.J.: General Learning Press, 1971.

Kleck, R. E. Physical stigma and nonverbal cues emitted in face-to-face interaction. *Human Relations*, 1968, *21*, 19–28.

Kleck, R. E. Physical stigma and task oriented interactions. *Human Relations*, 1969, *22*, 53–60.

Kleck, R. E., et al. The effect of stigmatizing conditions on the use of personal space. *Psychological Reports*, 1968, *23*, 111–118.

Kleck, R. E., & Nuessle, W. Congruence between indicative and communicative functions of eye contact in interpersonal relations. *British Journal of Social and Clinical Psychology*, 1968, *7*, 241–246.

Kleck, R. E., Ono, H., & Hastorf, A. H. The effect of physical deviance upon face-to-face interaction. *Human Relations*, 1966, *19*, 425–436.

Kleinke, C. L. *First impressions: The psychology of encountering others.* Englewood Cliffs, N.J.: Prentice-Hall, 1975.

MacGregor, F. C. *Transformation and identity.* New York: New York Times Book Company, 1974.

Maddox, G. L., Back, R. W., & Liederman, V. T. Overweight as social deviance and disability. *Journal of Health and Social Behavior*, 1968, *9*, 287–298.

McArthur, L. Z., & Solomon, L. K. Perceptions of an aggressive encounter as a function of the victim's salience and the perceiver's arousal. *Journal of Personality and Social Psychology*, 1978, *36*, 1278–1290.

Snyder, M. L., Kleck, R. E., Strenta, A., & Mentzer, S. J. Avoidance of the handicapped: An attributional ambiguity analysis. *Journal of Personality and Social Psychology*, 1979, *37*, 2297–2306.

Snyder, M., & Swan, W. B., Jr. Behavioral confirmation in social interaction: From social perception to social reality. *Journal of Experimental Social Psychology*, 1978, *14*, 148–162.

Wright, B. *Physical disability: A psychological approach*. New York: Harper & Row, 1960.

Received December 6, 1979 ■