

BEAUTY IS TALENT: TASK EVALUATION AS A FUNCTION OF THE PERFORMER'S PHYSICAL ATTRACTIVENESS

DAVID LANDY¹ AND HAROLD SIGALL²

University of Rochester

Male college subjects read an essay that supposedly had been written by a college freshman co-ed. They then evaluated the quality of the essay and the ability of its writer on several dimensions. By means of a photo attached to the essay, one third of the subjects were led to believe that the writer was physically attractive and one third that she was unattractive. The remaining subjects read the essay without any information about the writer's appearance. In addition, one half of the subjects read a version of the essay that was well written while the other subjects read a version that was poorly written. Significant main effects for essay quality and writer attractiveness were predicted and obtained. The subjects who read the good essay evaluated the writer and her work more favorably than the subjects who read the poor essay. The subjects also evaluated the writer and her work most favorably when she was attractive, least when she was unattractive, and intermediately when her appearance was unknown. The impact of the writer's attractiveness on the evaluation of her and her work was most pronounced when the "objective" quality of her work was relatively poor.

There is an increasing amount of research data attesting to the relative importance of physical attractiveness as a determinant and moderator of a wide variety of interpersonal processes: heterosexual liking (Berscheid, Dion, Walster, & Walster, 1971; Walster, Aronson, Abrahams, & Rottman, 1966), person perception (Sigall & Landy, 1973), persuasion (Mills & Aronson, 1965), peer popularity in young children (Dion & Berscheid, 1972), reactions to evaluations (Sigall & Aronson, 1969; Sigall, Page, & Brown, 1971), attributions of personal characteristics and future success (Clifford & Walster, in press; Dion, Berscheid, & Walster, 1972; Miller, 1970), and adult judgments of children's transgressions (Dion, 1972).

Several of these experiments have demonstrated that individuals tend to form impressions and make judgments about people on the basis of their physical attractiveness. For example, Miller (1970) found that subjects made more favorable attributions to good-

looking people than to unattractive people, and Dion et al. (1972) showed that college students of both sexes expected physically attractive men and women to possess more socially desirable traits (sensitivity, strength, modesty, etc.) than unattractive people. In addition, in Dion et al.'s study, students expected attractive people to have more good things in store for them in the future—more prestigious occupations and happier marriages—than unattractive people. Dion and Berscheid (1972) found a similar tendency in nursery school children in that attractive children, as judged by adults, were more popular with their peers and were seen to manifest less socially undesirable interpersonal behavior than their unattractive counterparts.

Clifford and Walster (in press) demonstrated that teachers expected physically attractive children to have greater academic potential and better social relationships with their peers than unattractive children. In this experiment, fifth-grade teachers examined a standardized report card containing identical objective information about a fifth-grade pupil. The photograph of either an attractive or unattractive boy or girl was pasted in a corner of the report card. The teachers' assessment of the pupil's IQ, his parents' attitudes

¹ Presently unaffiliated.

² The authors would like to thank James Gallant for his assistance.

Requests for reprints should be sent to Harold Sigall, who is now at the Department of Psychology, University of Maryland, College Park, Maryland 20742.

toward school, his future level of educational attainment, and his relationships with his peers were higher or more favorable when the pupil was attractive than when he was unattractive.

In the experiments described above, it was shown that individuals attribute more positive characteristics to and expect better performances from attractive people than unattractive people. Do individuals evaluate the actual performances of attractive people more positively than those of unattractive people? There is some evidence relevant to this question. In a correlational study of manipulative strategies, Singer (1964) examined the relationship between the physical attractiveness of female college students, based on ratings of photographs by faculty members, and the students' grade point average. A significant positive correlation ($r = .40$) was reported for (firstborn) females. However, given the possibility that there is a positive relationship between physical attractiveness and intelligence, attractive students may not merely have been the beneficiaries of perceptual bias. It is possible that attractive individuals are genetically endowed with greater ability or that they have environmental histories which produce greater intellectual capabilities, and that therefore they are better students.

Thus while there is experimental evidence that individuals expect attractive people to perform better, that is, have greater potential, and correlational evidence that individuals rate the work of attractive people more favorably, a direct causal relationship between physical attractiveness and performance evaluation has not yet been demonstrated.

The present study was designed to experimentally determine the effect of physical attractiveness on performance evaluation while controlling the quality of the task being evaluated and the evaluator's exposure to the performer. We expected that physical attractiveness would strongly influence the evaluation of an individual's performance on a given task even though the task was logically unrelated to appearance. The more attractive the performer, the more positive we expected the subjects' evaluations of her work to be. In addition to the attractiveness of the performer, we manipulated the quality of the

work being evaluated. While we expected the predicted relationship to be manifest for both high- and low-quality work, we wanted to explore the possibility that the impact of physical attractiveness on performance evaluation would vary with the ("objective") quality of the performance.

METHOD

Design

The subjects read a short essay and then evaluated its overall quality and gave their impressions of its writer. One half of the subjects read an essay that was well written while the other subjects read an essay that was poorly written. In addition, one third of the subjects were led to believe that the essay had been written by a physically attractive college co-ed, another third of the subjects were led to believe it had been written by an unattractive co-ed, and the remaining third—the control subjects—read either the well-written or the poorly written essay without being exposed to the attractiveness manipulation. This procedure yielded a 2×3 factorial design with two levels of essay quality (good and poor) and three levels of physical attractiveness (attractive, unattractive, and control).

Subjects

The subjects were 60 male undergraduates at the University of Rochester who were recruited from an introductory psychology course.

Essay Quality

Two standard untitled essays about the role and effects of television in society were prepared. Each essay was about 700 words in length and discussed similar issues, such as the effects of televised violence on children. However, one essay was well written, grammatically correct, organized, and clear in its presentation of ideas. The other essay was poorly written, contained numerous clichés and errors in usage, was disorganized, and simplistic in its presentation of ideas. Pretesting with a sample of 30 undergraduates demonstrated that the "good" essay was indeed viewed as a better essay than the "poor" essay. One half of the pretest sample read the good essay, and the rest read the poor essay. They then rated the essay's "general quality" on a 9-point scale labeled poor (1) and good (9) at the end points. The mean rating given to the good essay was 6.47 ($SD = 1.36$), while the mean rating given the poor essay was 4.40 ($SD = 1.76$). These means were significantly different ($F = 10.40$, $df = 1/28$, $p < .005$).

Physical Attractiveness of the Writer

Two college yearbook photographs of female students were selected for use in the experiment. One of the photographs was of a physically attractive woman, and one was of a physically unattractive woman. The selection of photographs was based on

the judgments of six male graduate students who were asked to rank order, in terms of physical attractiveness, a set of 15 facial photos of women from a recent college yearbook. The photo that had the highest mean ranking and the photo that had the lowest mean ranking were selected. So that six subjects could be tested at once, two copies of each photo were mounted on separate file cards on which a fictitious name, Marilyn Thomas, was typed along with a bogus standard description of the girl pictured. This information was the same on each card and described the writer as a freshman college student from Ohio whose father was a businessman, whose mother was a housewife, and who had two brothers and a sister. Her hobbies were identified as horseback riding and reading. All of the cards, including two cards that did not have photos mounted on them, contained the same background information. Each of these cards was attached to a copy of the essays by means of a paper clip. This was done in such a way that in order for the essay to be read, the card would have to be moved, presumably resulting in the readers' attending to the information provided, that is, the photograph on those cards having one, and the standard background information about the writer.

Procedure

Two to six subjects were scheduled to report for the experiment during any given experimental session. When the subjects arrived at the waiting room, they were greeted by the experimenter who explained that the experiment dealt with "social judgment, that is, how people make assessments of others." He told the subjects that he wanted them each to read and judge one of a number of essays that had been submitted in a freshman English class. The experimenter further explained that the instructor of this class had asked his students to prepare as an exercise an essay for submission to a contest being run by a local television station. Thus, each essay was turned in to the instructor along with some descriptive background about its author. These essays were then used by the experimenter to conduct the research in social judgment.

Following these introductory remarks, the experimenter ushered each subject into a separate cubicle where he randomly assigned him to one of the experimental conditions. He then handed each individual subject one of the sets of materials—a writer background information card attached to an essay—corresponding to the condition to which he had been assigned. There were 10 subjects in each of the six experimental conditions.

Dependent Measure

When each subject had finished reading the essay, the experimenter asked him to evaluate the essay by filling out a "judgment form." This consisted of eight rating scales. The subjects evaluated the essay they had read on four dimensions by circling the appropriate number on each of the 9-point rating

scales labeled at the end points, which followed each dimension—creativity: low (1)–high (9); ideas: dull (1)–interesting (9); style: poor (1)–good (9); and general quality: poor (1)–good (9). The subjects rated their impressions of the writer of the essay they had read on the following four dimensions—intelligence: low (1)–high (9); sensitivity: low (1)–high (9); talent: low (1)–high (9); and overall ability: low (1)–high (9).

When each of the subjects had completed filling out the judgment form, the experimenter asked them to recongregate in the waiting room. He then interviewed them. After assuring himself that the subjects were not suspicious, he informed them about the true nature of the experiment.

RESULTS AND DISCUSSION

Two measures of performance evaluation were derived from the subjects' responses on the judgment form. The first was simply their rating of the general quality of the essay as given by their reactions to a single rating scale. The second was an essay evaluation score consisting of the sum of the ratings of the essay on three dimensions: creativity, ideas, and style.³

Table 1 presents the means, standard deviations, and analysis of variance of ratings of the general quality of the essay. Table 2 presents the means, standard deviations, and analysis of variance of the essay evaluation scores. From an examination of these tables, it is clear that the less physically attractive the writer, the lower were the subjects' evaluations of her work. The main effect for writer attractiveness was significant on both the general quality measure ($F = 6.26$, $df = 2/54$, $p < .01$) and the essay evaluation measure ($F = 5.34$, $df = 2/54$, $p < .01$).

The highly significant main effect for essay quality on both measures of performance evaluation indicates that, as intended, one essay was perceived to be of higher quality than the other. This essentially constituted a check on the manipulation of essay quality.

The overall (2×3) interaction between essay quality and writer attractiveness was not significant on either measure of perform-

³ The intercorrelations of the subjects' ratings on these three dimensions ranged from .34 (creativity and style) to .65 (creativity and ideas). The mean ratings on these separate dimensions for subjects in each of the experimental conditions followed the same pattern as the means of the essay evaluation scores.

TABLE 1
RATINGS OF THE GENERAL QUALITY OF THE
ESSAY FOR SUBJECTS IN EACH OF THE
EXPERIMENTAL CONDITIONS

Means and standard deviations*				
Essay quality	Writer's physical attractiveness			Total <i>M</i>
	Attractive	Control	Unattractive	
Good				
<i>M</i>	6.7	6.6	5.9	6.4
<i>SD</i>	1.57	1.35	1.60	
Poor				
<i>M</i>	5.2	4.7	2.7	4.2
<i>SD</i>	1.55	1.95	1.34	
Total <i>M</i>	6.0	5.5	4.3	
Analysis of variance				
Source	<i>df</i>	<i>MS</i>	<i>F</i>	
Essay quality (A)	1	72.600	29.43**	
Writer attractiveness (B)	2	15.450	6.26*	
A × B	2	3.950	1.60	
Error	54	2.467		

* The higher the number, the better the quality; $n = 10$ in each condition.

* $p < .01$.

** $p < .001$.

TABLE 2
ESSAY EVALUATION SCORES FOR SUBJECTS IN EACH
OF THE EXPERIMENTAL CONDITIONS

Means and standard deviations*				
Essay quality	Writer's physical attractiveness			Total <i>M</i>
	Attractive	Control	Unattractive	
Good				
<i>M</i>	17.9	17.9	15.5	17.1
<i>SD</i>	4.82	3.60	4.70	
Poor				
<i>M</i>	14.9	13.4	8.7	12.3
<i>SD</i>	3.31	5.99	3.68	
Total <i>M</i>	16.4	15.6	12.1	
Analysis of variance				
Source	<i>df</i>	<i>MS</i>	<i>F</i>	
Essay quality (A)	1	340.817	17.24**	
Writer's attractiveness (B)	2	105.517	5.34*	
A × B	2	18.317	<1	
Error	54	19.772		

* Sum of ratings for creativity, ideas, and style; the higher the score, the more favorable the evaluation; $n = 10$ in all conditions.

* $p < .01$.

** $p < .001$.

ance evaluation. However, since the control condition in this experiment represented a base-line condition rather than an intermediate level of writer attractiveness, it makes sense to examine the 2×2 interactions, control conditions deleted, among the experimental conditions. On the whole these interactions suggest that physical attractiveness may have had greater impact when the quality of the work was poor than when it was good. With respect to general quality (Table 1), the 2×2 interaction produced an F of 2.93, ($df = 1/54$, $p < .10$). Furthermore, a simple comparison between the attractive-good (6.7) and the unattractive-good (5.9) conditions yielded a statistically nonsignificant finding ($F = 1.30$, $df = 1/54$), while the difference between the attractive-poor (5.2) and the unattractive-poor (2.7) conditions was highly significant ($F = 12.65$, $df = 1/54$, $p < .001$). The 2×2 interaction for essay evaluation scores produced equivocal results ($F = 1.83$, $df = 1/54$). However, the simple comparisons showed that the difference between the attractive-good (17.9) and unattractive-good (15.5) conditions was not significant ($F = 1.46$, $df = 1/54$), while the difference between the attractive-poor (14.9) and unattractive-poor (8.7) conditions was significant ($F = 9.71$, $df = 1/54$, $p < .005$).

A consideration of these findings permits some interesting speculation. One possible implication is that if someone's work is competent, personal characteristics are less subject to influence evaluations of that work than when the quality of the work is relatively poor. Thus, if you are ugly you are not discriminated against a great deal as long as your performance is impressive. However, should performance be below par, attractiveness matters: You may be able to get away with inferior work if you are beautiful. Furthermore, it should be noted that the standard deviations for the control group were consistently higher for the poor essay than for the good essay. This may indicate that the poor essay was a more ambiguous stimulus than the good essay. Thus it is possible to interpret the results as an indication that physical attractiveness plays a larger role in the evaluation of ambiguous rather than poor performances.

As noted earlier, Clifford and Walster (in press) found that attractive people, at least children, generate expectancies that they will perform well. Perhaps this expectancy leads evaluators to give physically attractive performers the benefit of the doubt when performance is substandard or of ambiguous quality. While admittedly speculative, this view is consistent with other findings concerned with discrimination based on "irrelevant" characteristics. For example, Rosenblatt (1967) presented evidence that in major league baseball blacks were not discriminated against much if they were good hitters; among players whose performances were relatively poor, however, whites were given greater opportunity, and blacks were victims of discrimination.

In addition to the measures of performance evaluation, the present authors collected data on the impact of essay quality and writer attractiveness on the subjects' impressions of the writer. There were two such measures: (a) a rating of the writer's overall ability indicated by the subjects' responses on a single rating scale and (b) a writer impression score consisting of the sum of the subjects' ratings of the writer on three dimensions—intelligence, sensitivity, and talent.⁴

Tables 3 and 4 present the means, standard deviations, and analysis of variance of the subjects' ratings of the writer's overall ability and the writer impression scores, respectively. The subjects rated the writer most favorably when she was physically attractive, least favorably when she was unattractive, and intermediately when no information about her appearance was given. This physical attractiveness main effect was manifested on both the overall ability measure ($F = 4.18$, $df = 2/54$, $p < .025$) and the writer impression measure ($F = 5.38$, $df = 2/54$, $p < .01$). Not surprisingly, the subjects also rated the writer more favorably in the good essay conditions than in the poor essay conditions. This produced significant main effects for essay quality

⁴ The intercorrelations of the subjects' ratings on these three dimensions ranged from .47 (sensitivity and talent) to .54 (sensitivity and intelligence). The mean rating on these separate dimensions for subjects in each of the experimental conditions followed the same pattern as the means of the writer impression scores.

TABLE 3
RATINGS OF THE ESSAY WRITER'S OVERALL ABILITY FOR SUBJECTS IN EACH OF THE EXPERIMENTAL CONDITIONS

Means and standard deviations ^a				
Essay quality	Writer's physical attractiveness			Total <i>M</i>
	Attractive	Control	Unattractive	
Good				
<i>M</i>	6.4	6.3	6.0	6.2
<i>SD</i>	1.07	1.06	1.94	
Poor				
<i>M</i>	5.7	4.7	3.4	4.6
<i>SD</i>	1.16	2.16	1.07	
Total <i>M</i>	6.5	5.6	4.7	

Analysis of variance			
Source	<i>df</i>	<i>MS</i>	<i>F</i>
Essay quality (A)	1	40.017	18.14**
Writer's attractiveness (B)	2	9.217	4.18*
A × B	2	4.517	2.05
Error	54	2.206	

^a The higher the number, the higher the ability; $n = 10$ in all conditions.
* $p < .025$.
** $p < .001$.

TABLE 4
WRITER IMPRESSION SCORES FOR SUBJECTS IN EACH OF THE EXPERIMENTAL CONDITIONS

Means and standard deviations ^a				
Essay quality	Writer's physical attractiveness			Total <i>M</i>
	Attractive	Control	Unattractive	
Good				
<i>M</i>	19.3	19.1	16.8	18.4
<i>SD</i>	4.06	2.73	4.52	
Poor				
<i>M</i>	17.2	14.8	11.4	14.5
<i>SD</i>	3.46	5.90	3.06	
Total <i>M</i>	18.3	17.0	14.1	

Analysis of variance			
Source	<i>df</i>	<i>MS</i>	<i>F</i>
Essay quality (A)	1	232.067	13.86**
Writer attractiveness (B)	2	90.117	5.38*
A × B	2	14.117	<1
Error	54	16.744	

^a Sum of impression ratings for intelligence, sensitivity, and talent; the higher the score, the more favorable the impression of the writer; $n = 10$ in all conditions.
* $p < .01$.
** $p < .001$.

on both measures of the subjects' impressions of the writer.

Looking at the 2×2 (Good-Bad \times Attractiveness) interactions, we find that with respect to the writer's overall ability (Table 3) the interaction was statistically significant ($F = 4.09$, $df = 1/54$, $p < .05$), and that the difference between the attractive-good (6.4) and unattractive-good (6.0) conditions was trivial, while the difference between the attractive-poor (5.7) and unattractive-poor (3.4) conditions was large ($F = 12.02$, $df = 1/54$, $p < .005$). For the data reported in Table 4, the 2×2 interaction was not statistically significant, although a simple comparison yielded a highly significant difference when the essay was poor and a nonsignificant difference when it was good. These findings, of course, are congruent with the speculation offered above.

The results of this experiment indicate that the physical attractiveness of an individual performing a given task affects the manner in which people evaluate both the performance and the performer. This was so even though the task performance being evaluated was completely unrelated to the physical attractiveness of the performer. Thus, physical appearance not only affects the way in which others react to a person, it also affects the way in which they react to that person's accomplishments.

As has been previously pointed out (Berscheid & Walster, 1972; Clifford & Walster, in press; Sigall & Landy, 1973), physical attractiveness forms one of the more pervasive bases for interpersonal discrimination. To the extent that we become aware of our tendency to evaluate the work of unattractive individuals less favorably than that of attractive individuals, we are, at least, able to make efforts to alter this tendency and the inequities it must produce.

It should be noted that in this study male subjects responded to female stimulus persons. Whether or not the relationships found would hold for other sex combinations is an open question. Clearly the results should not be too hastily generalized. On the other hand, there is evidence (e.g., Byrne, London, & Reeves,

1968; Dion et al., 1972; Sigall & Landy, 1973; Walster et al., 1966) that both men and women are rather susceptible to having their judgments influenced by the physical attractiveness of others.

REFERENCES

- BERSCHEID, E., DION, K., WALSTER, E., & WALSTER, G. W. Physical attractiveness and dating choice: A test of the matching hypothesis. *Journal of Experimental Social Psychology*, 1971, 7, 173-179.
- BERSCHEID, E., & WALSTER, E. Beauty and the best. *Psychology Today*, 1972, 5(10), 42-46.
- BYRNE, D., LONDON, O., & REEVES, K. The effects of physical attractiveness, sex, and attitude similarity on interpersonal attraction. *Journal of Personality*, 1968, 36, 259-271.
- CLIFFORD, M. M., & WALSTER, E. The effects of physical attractiveness on teacher expectation. *Sociology of Education*, in press.
- DION, K. Physical attractiveness and evaluation of children's transgressions. *Journal of Personality and Social Psychology*, 1972, 24, 207-213.
- DION, K., & BERSCHEID, E. Physical attractiveness and social perception of peers in pre-school children. Unpublished manuscript, University of Minnesota, 1972.
- DION, K., BERSCHEID, E., & WALSTER, E. What is beautiful is good. *Journal of Personality and Social Psychology*, 1972, 24, 285-290.
- MILLER, A. G. Role of attractiveness in impression formation. *Psychonomic Science*, 1970, 19, 241-243.
- MILLS, J., & ARONSON, E. Opinion change as a function of the communicator's attractiveness and desire to influence. *Journal of Personality and Social Psychology*, 1965, 1, 173-177.
- ROSENBLATT, A. Negroes in baseball: The failure of success. *Transaction*, 1967, 4(9), 51-53.
- SIGALL, H., & ARONSON, E. Liking for an evaluator as a function of her physical attractiveness and nature of the evaluations. *Journal of Experimental Social Psychology*, 1969, 5, 93-100.
- SIGALL, H., & LANDY, D. Radiating beauty: The effects of having a physically attractive partner on person perception. *Journal of Personality and Social Psychology*, 1973, 28, 218-224.
- SIGALL, H., PAGE, R., & BROWN, A. C. Effort expenditure as a function of evaluation and evaluator attractiveness. *Representative Research in Social Psychology*, 1971, 2, 19-25.
- SINGER, J. E. The use of manipulative strategies: Machiavellianism and attractiveness. *Sociometry*, 1964, 27, 128-150.
- WALSTER, E., ARONSON, V., ABRAHAMS, D., & ROTTMAN, L. The importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology*, 1966, 4, 508-516.

(Received July 3, 1972)