A series of studies was performed to investigate the effect of context upon evaluations made of personality-trait adjectives. Ss were presented sets of 3 adjectives which varied in favorableness as indicated by normative data and were then asked to evaluate 1 (test) adjective in each set under a variety of instructional conditions. Results indicated that whenever Ss were asked to rate the 3 adjectives in each set as a collective before rating test adjectives, evaluations of test adjectives increased with the favorableness of the context adjectives accompanying them. This effect occurred regardless of whether Ss were told that the adjectives in the collective described a single person or whether they were related in any way other than through their physical juxtaposition; it was only necessary that Ss form an impression of the collective as a whole before evaluating the test adjective. Results were interpreted as inconsistent with theoretical formulations of Anderson and Lampel (1965) and Osgood and Tannenbaum (1955). It was tentatively hypothesized that when Ss are asked to evaluate a group of stimuli, the range of stimuli included in each category along the scale used in making judgments, and hence the size of each scale category, increases with the dispersion of stimuli contained in the group.

Anderson and Lampel (1965) reported that when sets of trait descriptions (adjectives) were not ascribed to a hypothetical person, the rating of any particular trait in a set was independent of the quality of the adjectives with which it was presented (their context). However, when subjects were told to assume that the adjectives in each set described a person, the evaluation of any given adjective increased with the normative favorableness of the other members of the set. This latter effect was replicated in a later study (Anderson, 1966).

There are several possible explanations of the context effect described above. For example, the "meaning" of a trait adjective may change when it is associated with others. This point of view has been elaborated by Asch (1946). A shift in evaluative meaning (favorableness) may result in part from an attempt to reduce incongruity or imbalance produced by differences in the favorableness of the adjectives being associated (cf. Osgood, Suci, & Tannenbaum, 1957; Osgood & Tannenbaum, 1955). An explanation that does not require the assumption of a change in meaning has been offered by Anderson and Lampel (1965), who hypothesized that under the conditions of their study the evaluation of a trait adjective, if it is used to describe a person, can be represented as a weighted average of the evaluation of the attribute presented in isolation and the overall evaluation of the person possessing the attribute. If either of the above explanations is correct, the effects of context on ratings of adjectives should occur when the adjective sets describe a hypothetical person, but would be less likely to occur when the adjectives are not associated in a single stimulus.

On the other hand, it is conceivable that a general response process occurs when evaluating a set of more or less diverse stimuli as a collective that affects subsequent ratings of the individual stimuli. The precise nature of this process is unclear. However, if this were the case, a context effect might occur even when adjective sets do not describe a
single object, provided subjects are asked to evaluate the adjectives in each set as a group before they rate a particular adjective in the set.

A set of experiments was performed to investigate this possibility. In the first experiment an attempt was made to replicate the Anderson and Lampel (1965) results using a somewhat different experimental procedure. In the second and third experiments test adjectives were evaluated under conditions in which adjectives were rated as a collective, but in other ways were dissociated from one another. Whether the context effect would occur if subjects were instructed simply to form an impression of the adjectives presented as a group but were not required to rate the group on the basis of this impression was also investigated.

**Method**

Three experiments were performed. The general procedure used in each was similar. Subjects (introductory psychology students) were presented sets of three adjectives on slides. They were asked to estimate the likeableness of one of the adjectives on each slide and in some instances the likeableness of the set of adjectives as a group under various instructional conditions which are described in detail below.

**Preparation of Materials**

**Experiment I.** Twelve adjectives, three at each of four levels of favorableness, were selected from a list of 555 personality-trait adjectives compiled by Anderson (1964) for which normative ratings along a scale from 0 to 6 were available. Three adjectives (INTELLIGENT, LOYAL, and THOUGHTFUL) were designated as extremely favorable (H); three others (PERSUASIVE, QUICK, and SOFT-SPOKEN), as moderately favorable (M+); three (ARGUMENTATIVE, IMPUDENT, and FRIVOLOUS), as moderately unfavorable (M−); and three (VULGAR, INSOLENT, and MALICIOUS), as extremely unfavorable (L).

The mean favorableness of the adjectives selected at these four levels was 5.4, 3.8, 2.3, and 1.7, respectively. The first adjective in each set of three listed above was used as a test adjective, while the remaining two were designated as context adjectives. Each of the four test adjectives was then paired with each of the four pairs of context adjectives to form 16 different three-adjective sets (HHH, HHM+, HHM−, etc.).

In addition, two groups of 22 three-adjective filler sets were selected. Sets in one group were comprised of adjectives on the Anderson (1964) list designated as highly favorable; sets in the other group contained only highly unfavorable adjectives. The effects of between-group differences in the favorableness of these sets are not reported in this paper since they are not relevant to the primary question under consideration.

Each subject was presented 38 three-adjective sets: six filler sets followed by 16 test sets alternated with the remaining 16 sets from the filler group. The 16 sets of test adjectives were ordered randomly, with the restriction that at least five sets of adjectives separated sets containing the same test adjective. Each set of three adjectives was projected simultaneously on a screen. Adjectives were arranged in a column. The position of the test adjective in each set of three was systematically varied to minimize order effects.

The present study differed methodologically from the Anderson and Lampel (1965) experiment in the number of favorableness levels of test adjectives presented and in the frequency with which each subject was exposed to the same adjective. The design was more similar to the later study by Anderson (1966). While the number of adjectives selected for this study was very limited, it was assumed that a successful replication of the Anderson and Lampel findings would indicate that the set of stimuli selected was appropriate for further investigation of context effects.

**Experiments II and III.** In these two experiments, the filler sets of stimuli used in Experiment I were eliminated, and a second stimulus replication was added. Sixteen sets were constructed in the manner described in Experiment I. The extremely favorable (H) adjectives selected were HONEST, DEPENDABLE, and GOOD-NATURED (M = 5.4); the moderately favorable (M+) adjectives were DISCIPLINED, PERSISTENT, and OBJECTIVE (M = 3.6); the moderately unfavorable (M−) adjectives were GULLIBLE, UNINQUISITIVE, and TIMID (M = 2.2); and the extremely unfavorable (L) adjectives were INOBLIGANT, INSINCERE, and SELFISH (M = .8). The first in each set of three was used as the test adjective.

A sequence of 32 sets of adjectives was constructed by alternating the above sets of adjectives (Group B) with the 16 test sets used in Experiment I (Group A), counterbalancing order of sets within set groups. These 32 sets were preceded by six practice sets which contained adjectives spanning the entire favorableness continuum as indicated by normative data.

**Instructional Conditions**

**Experiment I.** Thirty-three subjects, 16 who received unfavorable filler sets and 17 who received favorable filler, were exposed to each of two instructional conditions, labeled P and W to conform to Anderson and Lampel's (1965) notation. These conditions are similar to those used in the earlier study except that subjects were run in groups ranging in size from 8 to 10 rather than individually, and subjects made responses in writing rather than orally.

Condition P was intended to give subjects the impression that the adjectives presented in each set described a person. The following instructions were given:
We would like to determine how persons form impressions of others and how they interpret information about them. Each of the following slides will show three adjectives that were used to describe a certain person on the basis of an informal interview. When I present a slide, form an impression of what the person described would be like and then (by placing a check along the appropriate scale) indicate how well you would like this person. After you have done this, I will ask you to judge how well you would like a particular trait of the person that I will point to and to record this judgment.

In Condition W, which was intended to dissociate the three adjectives in each set from one another, the following instructions were given:

We would like to determine how persons evaluate personality traits. On each of the following slides there will be three adjectives. These adjectives are not intended to apply to any given person. In fact, as you will see, it would often be very difficult to think of them as describing a single individual. When I present each slide, estimate to yourself how much you like each of the three traits shown. Then, when I point to one of the traits, indicate how much you like it by placing a check on the first scale on your answer sheet.

Under each condition, judgments of persons and adjectives were made along similar 21-point scales ranging from 0 (dislike very much) to 20 (like very much).

Experiment II. In this experiment, instructional conditions P and W were repeated. Two additional conditions were run in which an attempt was made to dissociate the adjectives contained in each set from one another while they nevertheless would continue to be evaluated as a collective. The number of subjects run under each condition ranged from 14-17.

In one condition, W, the adjectives were not associated with a person. The instructions paralleled those used in Condition W; that is, subjects were told that the adjectives shown on each slide were not intended to apply to any given person and that, in fact, it would often be very difficult to think of them as describing a single individual. However, they were told that when each slide was presented they should "try to arrive at a single estimate of how well [they] would like the traits shown considered in combination" and to record this estimate on their answer sheets. They were then asked to estimate their liking for the test adjective contained in the set.

In the second condition, Ps, an attempt was made to dissociate the adjectives contained in each set, even though they were used to describe a single person:

We would like to determine how persons form impressions of others and interpret information about them. On each of the slides I will show you there are three adjectives that were used to describe a certain person on the basis of informal interviews. Some of the persons described have what are often called "stable" or well integrated personalities; others, however, have what are known as disorganized personalities; these persons are highly changeable and often shift abruptly from one type of behavior to another. The sets of slides I will show you will describe several persons of this disorganized type. When I present a slide, form an impression of what the person described would be like... etc.

Experiment III. In this study, Conditions P and W were repeated, and three conditions were added. Instructions in two of these conditions, P, and Wo, were similar to those given in Conditions P and W, except that, instead of asking subjects to rate the collective along a scale, they were simply told to form an impression of the person (or of "the adjectives considered as a group") and then, when the experimenter pointed to the test adjective, to record their liking for the trait described by this adjective.

In the fifth condition, G, subjects were told to assume that each adjective described a different person. They were asked to form an impression of the group of persons and to record the estimate on the first scale provided. Then, when the experimenter pointed to a particular adjective, they were told to estimate how much they would like that particular person in the group.

Twenty-seven subjects were run under Condition P, 27 under Condition W, 12 under Condition G, 16 under condition P, and 16 under Condition W.

After ratings had been completed, subjects were asked to complete a short questionnaire pertaining to the procedure they used when making ratings. Under Conditions W and W, subjects were asked specifically whether they thought of a person when forming an impression of a set of adjectives. Under all conditions, subjects were asked if they found that they had had to evaluate each test adjective in relation to those accompanying it or if they evaluated the test adjective by itself.

Of 34 subjects run under Conditions W and W, who reported using a specific procedure for rating test adjectives, only 6 reported rating adjectives in relation to the context adjectives rather than independently. Of 37 subjects run under Conditions P and W, who reported using specific procedures, 20 reported rating adjectives in relation to the context adjectives. The difference between these proportions is significant ($x^2 = 5.94, df = 1, p < .02$). Under Condition G, only 2 of 12 subjects reported rating test adjectives in relation to the context. Instructions that adjectives were unrelated to one another therefore appear to have been effective.

Results

Experiment I

Two instructional conditions were run: P (in which adjectives in each set were said to
describe a single person), and W (in which each adjective in the set was evaluated individually). An extension of a Lindquist (1953) Type VI analysis of variance was performed on favorableness ratings of test adjectives as a function of the normative favorableness level of the test adjective (T), the favorableness of the context adjectives (C), instructional conditions (I), and the favorableness of filler sets. Results involving the last variable above are not reported in light of their peripheral relevance to the question on which this paper is primarily focused. In only one instance did this variable interact significantly with other independent variables under consideration (T and C).

The analysis produced significant main effects of I \((F = 4.56, df = 1/62, p < .05)\), C \((F = 17.63, df = 3/186, p < .01)\), and T \((F = 195.04, df = 3/182, p < .01)\), and significant interactions of C X I \((F = 17.63, df = 3/186, p < .01)\) and C X T \((F = 1.94, df = 9/558, p < .05)\).

The C X I interaction, described in Figure 1, indicates that ratings of test adjectives were displayed toward the favorableness of the context adjectives only under Condition P, thus replicating the results obtained by Anderson and Lampel (1965). This effect was not contingent upon the favorableness of the filler sets interspersed among test sets.

**Evaluation of persons.** Evaluations of the persons described by the sets presented under Condition P were analyzed as a function of the favorableness of the test adjective (T), the favorableness of context adjectives (C), and the favorableness of filler sets. A significant interaction of T X C occurred \((F = 7.76, df = 9/93, p < .001)\). The nature of this interaction is shown in Figure 2. Since only one test trait and one set of context traits at each level of favorableness were used in this study, the generality of these results is limited. However, these data nevertheless question the adequacy of simple additive and averaging formulations of attitude formation considered by other authors (e.g., Anderson, 1962, 1964).

**Experiment II**

In this study, four instructional conditions were used: P, W, P, (in which adjectives allegedly described persons with “disorganized personalities”), and W (in which adjective...
sets were not said to describe a person, but yet were evaluated as a collective). An analysis of variance was performed as a function of the favorableness of test adjectives (T), the favorableness of context adjectives (C), stimulus replications (R), and instructional conditions (I). A summary of this analysis is shown in Table 1.

Of primary relevance to the issue under consideration is the C × I interaction. The nature of this interaction is described in Figure 3. No effect of context occurred under Condition W, in which adjectives were not considered as a collective, while the context effect was greatest under Condition P, in which adjective sets were said to describe a person.

The effect of C under Condition W, collapsed over stimulus replications, was compared with the effect of C under each of the other three instructional conditions. A significant simple interaction of C and I was assumed to indicate a tendency to displace the rating of the test adjective toward the favorableness of the context adjectives under the condition compared with W. This interaction was significant in comparing Condition W with Conditions P (F = 11.60, \( p < .01 \)),
W (F = 9.63, p < .01), and P (F = 3.17, p < .025).

To determine whether the context effects under Conditions W and P were similar in magnitude to the context effect under Condition P, additional simple-effects analyses were performed. The C × I interaction was not significant in comparing P with W (F = .19), but reached significance in comparing P with P (F = 2.86, p < .05). This significance is attributable to the fact that under Condition P, adjectives were evaluated no more favorably when the context was highly favorable than when it was moderately favorable.

To summarize, Condition W, in which adjectives were not ascribed to a person, produced as strong a context effect as Condition P, in which adjectives described a single individual. Condition P, in which adjectives described a person with a disorganized personality, produced a displacement of the test-adjective rating toward the context, but this effect was not as pronounced as under Condition P.

While no hypotheses were made concerning the interactive effects of T and C, the fact that this interaction was significant in the analysis of Experiment II (and of Experiment I) is worth noting. A more detailed discussion of the implications of this interaction will be deferred to a later section. Let it suffice to point out here that the interaction runs counter to the explanation of the context effect proposed by Anderson and Lampel (1965).

Evaluation of collectives. Supplementary analyses were performed on evaluations made of collectives under Conditions P, P, and W. These analyses seemed of particular interest in light of the significant T × C interaction found during Experiment I. No effects involving instruction conditions were significant. However, the T × C interaction was significant, even when each stimulus replication was analyzed separately (for Group A data, F = 6.98, df = 9/135, p < .01; for Group B data, F = 5.29, df = 9/135, p < .01). These results, which were not contingent upon instructional conditions, question the validity of the hypothesis that ratings of collectives are a simple function of either the sum or the average of the component values.

Experiment III

In this experiment, Conditions P and W were repeated. Two additional conditions, P and W, were similar to P and W, except that collective evaluations were not recorded. In a fifth condition, G, each adjective was said to describe a different person.

An analysis of variance of ratings of test adjectives was performed as a function of T, C, I, and R, as defined in Experiment II. A summary of these data is shown in Table 1. While the main effect of context was highly significant, an interaction of context and instructional condition did not occur (F = 1.13).

The nature of the context effect under each instructional condition is shown more clearly in Figure 4. Ratings were somewhat erratic, but generally increased with favorableness of the context. No consistent differences between instructional conditions are apparent. Under Condition G, adjectives presented in highly unfavorable contexts were rated higher than adjectives presented in moderately unfavorable contexts. This tendency, which was not strong enough to produce a significant C × I interaction, was the only indication that the effect of context was different under instructions that adjectives described different stimuli than under instructions that they described a single stimulus.

![Fig. 4. Evaluation of test adjectives as a function of instructional conditions and favorableness of context adjectives—Experiment II.](image-url)
Supplementary analyses. It was speculated that the similarity of context effects under Conditions P and W_c might be due to the fact that subjects under Condition W_c, when rating adjectives as a collective, associated them with a single person despite instructions to consider them as independent. Subsequent questioning of subjects run under this condition indicated that 15 subjects reported considering adjectives independently of persons, while 9 reported ascribing them to a single person. Analyses of ratings made under Condition W_c by subjects who reported using each procedure yielded a nonsignificant interaction of context favorableness, and rating procedure was not significant \( F = 1.17, p > .25 \). It seems reasonable to conclude that the effect of context on ratings of test adjectives does not depend upon whether these adjectives describe a person.

DISCUSSION

The series of experiments reported here eliminated a number of possible explanations of the apparent displacement of adjective ratings toward the context in which they are presented. While a necessary condition for the effect is apparently that an impression of the set of adjectives as a collective be formed prior to rating the test adjective, it is not necessary that these adjectives be ascribed to a single stimulus object or be related in any way other than through their presence on the same slide. The effect cannot be attributed to a response set created by evaluating the collective on a scale similar to that used in rating the test adjective. While the possibility cannot be completely discounted that the meaning of the test adjective changes as a result of its presentation in a collective, the occurrence of context effects under Conditions W_c and G decreases the likeliness of this explanation.

Anderson and Lampel (1965) hypothesized that the impression of a person is an average of the values of the separate characteristics ascribed to him, and that the rating of the test adjective is a weighted average of the subject’s liking for the test adjective in isolation and his liking for the person to whom it is ascribed. A necessary condition for the support of this hypothesis is that the interaction of test-adjective favorableness and context favorableness be zero. However, data collected in Experiments I–III indicate that in each analysis the interaction was significant. Some caution should be taken in interpreting this interaction, since ratings may not have been made along an equal interval scale. Also, the limited number of adjectives used in these experiments prevents much generality from being attached to these findings. Nevertheless, the occurrence of an interaction involving any set of adjectives questions the validity of Anderson and Lampel’s hypothesis.

The incongruity formulation of Osgood and Tannenbaum (1955) predicts that the evaluation of an attribute will be displaced toward the evaluation of the object possessing the attribute, and that the magnitude of this displacement will increase with the difference between the evaluation of the object and the evaluation of the attribute in isolation. The theory is predicated on the assumption that the evaluations in question are in fact incongruous and therefore produce pressure to reduce the incongruity. There is no reason to believe that such an incongruity exists when adjectives are unrelated to one another, as in Condition W_c, or when they are ascribed to different persons, as in Condition G; the fact that context effects occur under these conditions suggests that the Osgood and Tannenbaum formulation is inapplicable.

It is possible that the context effect is attributable to a general response process (as opposed to a “perceptual” process) that is affected by rating the collective before rating the component. When rating a collective, subjects are required to represent more or less diverse stimuli by a single value or to place them in a single scale category. The range of these stimuli may affect the size of the category to which the collective is assigned; the range of stimuli assigned to other available categories is similarly affected. Specifically, in making a single rating of a collective of stimuli with diverse values, the size of each rating-scale category may be expanded; that is, the range of stimuli classified within each category may be increased. This would result in less extreme ratings of stimuli than
would be the case if the scale-unit size were small.

If category size is a positive function of the dispersion of stimulus values in each collective, the effect of context in the experiments reported here would generally be explained. The range of stimuli included in each scale category would be less when evaluating a collective containing an unfavorable test adjective and two equally unfavorable context adjectives than when evaluating a collective containing the same test adjective and two highly favorable context adjectives. The test adjective, evaluated along the same scale as the collective, would therefore be rated a greater number of categories below the origin in the former case than in the latter. Similarly, a highly favorable test adjective should be rated a greater number of categories above the origin when it is presented with two highly favorable context adjectives than when it is accompanied by two unfavorable adjectives.

The significant interaction of test-adjective favorableness and context favorableness in analyses of collective ratings suggests that a simple additive or averaging formulation of attitude formation (Anderson, 1962, 1964; Fishbein, 1963) may be inappropriate. If adjectives have simple additive effects upon impressions formed of social stimuli, and if equal interval scaling is assumed, the effect of each adjective upon evaluations of fictitious persons should not depend upon the quality of the two traits accompanying it. Possibly information that is incongruous with other information presented is discounted. Anderson and Jacobson (1965) found evidence that discounting occurs unless subjects are specifically instructed to disregard traits that appear inconsistent with other information available. A second interpretation, which would be consistent with the general hypothesis that additive and averaging models can effectively describe judgments made of social stimuli, is that the amount of additional information provided by each adjective presented varies with the quality of the adjectives accompanying it. For example, FRIENDLY may have relatively less effect upon judgments than would information that a VULGAR person is FRIENDLY if WARM and FRIENDLY are commonly associated, but VULGAR and FRIENDLY are not. Unfortunately, this interpretation does not seem applicable to Condition W, in which the Context × Test interaction was also apparent.

REFERENCES


(Received April 15, 1967)