The Effect of Rating Instructions on the Gain-Loss Ordering of Student Ratings of Teaching Performance

By

Shelly J. C. Turcotte

University of Manitoba

A Thesis

submitted to the Faculty of Graduate Studies

in Partial Fulfillment of the requirements for the

Degree of Masters of Arts

Running Head: Rating Instructions and Gain-Loss Ordering

THE EFFECT OF RATING INSTRUCTIONS ON THE GAIN-LOSS ORDERING OF STUDENT RATINGS OF TEACHING PERFORMANCE

BY

SHELLY J. C. TURCOTTE

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF ARTS

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my advisor, Dr. Les Leventhal, for his expert guidance, support, time and patience. His advice and calming influence were truly invaluable to me in the completion of this thesis.

I wish to thank my committee members, Dr. Steven Holborn and Dr. Lauran Sandals, for providing constructive suggestions and counsel under unusual time constraints. I would like to express my further appreciation to Dr. Holborn for supplying the encouragement and support that I needed to enter the Graduate program.

Finally, I am deeply indebted to my friend and mother, Mrs. Yvonne Turcotte, who typed countless drafts of the thesis within a few months and who demonstrated a remarkably high frustration threshold. I further wish to express my heartfelt gratitude to her for the other innumerable ways she has facilitated my progress.

ABSTRACT

In a laboratory analogue of a classroom, using videotaped lectures, the effects of initial teaching (good or poor lecture 1), final teaching (good or poor lecture 2) and instructions (consider only lecture 2, consider lecture 1 and 2) on final student ratings and liking were investigated. The relation of affect and self-esteem to ratings and liking was also examined. For students considering only lecture 2, lecture 1 had a negative effect (termed a negative primacy effect) and lecture 2 had a larger positive effect (termed a positive recency effect), supporting gain-loss theory. For students considering both lecture 1 and lecture 2, lecture 1 had a positive affect (termed a positive primacy effect) and lecture 2 had a larger positive effect (termed a positive recency effect), supporting reinforcement-affect theory. Evidence regarding the role of affect was inconclusive. Self-esteem was not related to ratings or liking. Results were discussed in terms of replacement versus addition of information as well as in terms of implications for previous research and teaching in the classroom. An alternative explanation involving contrast effects was also discussed.

The Effect of Rating Instructions on the Gain-Loss Ordering of Student Ratings of Teaching Performance

The relationship between the quality of lecturing in the classroom and student ratings of the instructor may be viewed as the effect of reinforcing and punishing stimuli on interpersonal attraction. Two models of interpersonal attraction are the reinforcement-affect model (Byrne, 1971; Clore & Byrne, 1974) and the gain-loss model (Aronson, 1969; Aronson & Linder, 1965). The reinforcement-affect model proposes that evaluative responses should be most positive towards stimuli associated with consistent reinforcement and most negative towards stimuli associated with consistent punishment. In contrast, the gain-loss model proposes that evaluative responses should be most positive towards stimuli associated with a gain in reinforcement and most negative towards stimuli associated with a loss in reinforcement. These models may be applied to end-of-course ratings if one assumes that, within a single course, lectures may be consistently good or consistently poor, or may gain or deteriorate in quality.

The effects of initial and final teaching performance on end-of-course student ratings were investigated by Leventhal, Turcotte, Abrami and Perry (Note 1). In a laboratory analogue of a classroom, initial lecture quality (good or poor lecture 1) and final lecture quality (good or poor lecture 2) were manipulated with prepared videotapes. Eight student rating measures were employed. Ratings of lec-

ture 2 were highest for the poor-good condition (poor lecture 1 and good lecture 2) followed in order by the good-good, poor-poor, and good-poor conditions. The results of the Leventhal et al. study suggest that lecturers whose teaching performance improves from poor to good receive higher student ratings than those who are consistently good. Further, lecturers whose teaching performance deteriorates from good to poor receive lower student ratings than those who are consistently poor. Leventhal et al. pointed out that Ramagli and Greenwood (Note 2), who addressed different issues, nevertheless included data which showed the same ordering of the four lecture sequences. Thus, the pattern of student ratings found by Leventhal et al. and Ramagli and Greenwood is consistent with the gain-loss model because a gain in lecture quality earned higher lecture 2 ratings than consistently good lecture quality and a loss in lecture quality earned lower lecture 2 ratings than consistently poor lecture quality.

The "gain-loss-ordering" (from highest to lowest: poor-good, good-good, poor-poor, and good-poor) found in the Leventhal et al. study was described by the authors in terms of primacy-recency effects. Lecture 2 quality had a large positive effect on ratings (termed a positive recency effect) while lecture 1 had a smaller negative effect on ratings (termed a negative primacy effect). The gain-loss ordering occurs when a negative primacy effect and a positive recency effect are present and the recency effect is larger than the primacy effect. From this perspective, the negative primacy effect is responsible for the poor-good "gain" teacher receiving

higher ratings than the good-good consistent teacher and for the poor-poor consistent teacher receiving higher ratings than the good--poor "loss" teacher. As explained in detail by Leventhal et al., primacy and recency effects describe any ordering.

Purposes of the Present Experiment

Leventhal et al. instructed students to evaluate only the final portion of the sequence (lecture 2). It is possible that the gain-loss ordering found may not have occurred with instructions to the students which did not emphasize lecture 2. Therefore, the main purpose of this study was to investigate the effect on the gain-loss ordering of instructions to consider only lecture 2 when evaluating the instructor as compared with instructions to consider both lecture 1 and lecture 2.

Both interpersonal attraction models have prompted research and theory which suggest that reinforcing and punishing stimuli influence evaluative responses by altering the affective state of the evaluator (e.g., Gouax, 1971; Mettee, 1973). Hence, a secondary purpose of this study was to examine the relation of student affective state and self-esteem to their ratings of the instructor. Another secondary purpose was to examine whether liking as measured in interpersonal attraction research would show the same results as ratings.

Importance of Instructions in Evaluating Sequences of Stimuli

Research on the effect of initial and later information on evaluation, i.e. primacy and recency effects, suggests that those effects are influenced by the nature of the instructions to the subjects. Primacy effects may be minimized by instructing subjects to respond

after each stimulus rather than at the end of the sequence (Bramel, 1969; Byrne, Lamberth, Palmer & London, 1969) and by interjecting some other activity or a period of time between each stimulus (Luchins, 1957; Rosenkrantz & Crockett, 1965).

Primacy effects may also be reduced by instructions which lead subjects to pay more attention to later information than initial information (Leach, 1974). Students in the Leventhal et al. study were expressly directed to evaluate the second half of the lecture sequence (lecture 2). If either the size or direction of the lecture 1 or lecture 2 effects differ with different instructions, student ratings may not show the gain-loss ordering. This is true because the ordering depends on a positive recency effect and a smaller negative primacy effect. For example, in the Leventhal et al. study, the effects of the initial lecture may have been small due to the students' attempts to disregard it. When students are instructed to consider both lectures, the primacy effect may be as large or larger than the recency effect, resulting in a different ordering. This study compared the effects on evaluating the instructor of both types of instructions: One group was instructed to consider only lecture 2 and another group was instructed to consider both lecture 1 and lecture 2. In this manner it was possible to determine if the gain-loss ordering occurs when instructions do not emphasize lecture 2.

Models of Interpersonal Attraction

The reinforcement-affect model of interpersonal attraction (Byrne & Clore, 1971; Clore & Byrne, 1974) holds that interpersonal events

can be classified as either reinforcing or punishing. Rewarding events elicit postive affect while punishing events elicit negative affect. The evaluation of any stimulus, including a person, is a function of the proportion of reinforcements and punishments associated with it. Evaluative responses are mediated by the positive affect accompanying reinforcement and the negative affect accompanying punishment. Any stimulus which is associated with reward or punishment and its accompanying positive or negative affect becomes a conditioned reinforcing or punishing stimulus which evokes the affective state. Stimuli producing positive affect are liked while stimuli evoking negative affect are disliked.[1]

The gain-loss model of interpersonal attraction (Aronson, 1969; Aronson & Linder, 1965; Mettee & Aronson, 1974) also proposes that reinforcing and punishing events produce affective states which determine liking. However, unlike the reinforcement-affect model, the gain-loss model maintains that sequence of positive and negative events is more important than quantity or proportion. Aronson and Linder (1965) had subjects overhear a series of evaluative remarks about themselves from another person (actually a confederate) which were either consistently positive, negative changing to positive, consistently negative, or positive changing to negative. Subjects' overall liking of the confederate was most favourable when the remarks changed from negative to positive, followed, in order, by the consistently positive, consistently negative, and positive to negative conditions. The greater liking for the poor-good evaluator than the good-good evaluator was termed a gain effect. Similarly, the lower

liking for the good-poor evaluator than the poor-poor evaluator was termed a loss effect. Aronson and Linder concluded that gain or loss in favourable feedback is more potent reward or punishment than invariant favourable or unfavourable feedback.

A major reason proposed by Aronson and Linder for the gain-loss effect involves the affective state of the subject. Invariant positive feedback produces positive affect and therefore liking of the stimulus person. A gain in favourableness of feedback (negative to positive) not only produces positive affect but also reduces negative affect produced by the original negative feedback. Reduction of negative affect serves as an added source of reward resulting in yet greater liking. Similarly, invariant negative feedback produces negative affect and therefore low liking. A loss in favourableness of feedback (positive to negative) not only produces negative affect but also reduces the positive affect produced by the original positive feedback. Reduction of positive affect serves as an added source of punishment resulting in yet less liking.

Subsequent to Aronson and Linder's 1965 study, research on the gain-loss effect in the areas of opinion change (Sigall & Aronson, 1967), non-verbal behaviour (Clore, Wiggins & Itkin, 1975), non-verbal immediacy (Coutts, Schneider & Montgomery, 1980) and the double-evaluator situation (Bersheid, Brothen & Graziano, 1976) provided partial support for the model. However, several studies have reported discrepant findings. Byrne and London (1966) found no significant difference in liking between a gain condition (negative-positive stimuli) and a loss condition (positive-negative stimuli). Taylor, Altman

and Sorrentino (1969) found that a negative-positive and a positive-positive sequence elicited equal liking and a positive-negative and a negative-negative sequence produced equal disliking. Gain effects have been more consistently found than loss effects (Mettee and Aronson, 1974). The results of studies by Hewitt (1972) and Tognoli and Keisner (1972) were consistent with the reinforcement model rather that with the gain-loss model: Persons associated with consistently positive reinforcement were liked most and persons associated with consistent punishment were liked least. [2]

A study by Clair and Snyder (1979) is particularly interesting because it examined gain-loss effects in student ratings of instruction. Clair and Snyder investigated the effects on ratings and academic performance of an instructor's feedback to students about their performance on a verbal learning task. Students were exposed to one of four instructor-delivered evaluative feedback conditions (uniformly positive, uniformly negative, negative to positive, or positive to negative). Students then listened to an audiotaped lecture, completed an exam on the lecture, and rated the instructor. Clair and Snyder's results supported the reinforcement-affect model. They found that ratings and performance were highest in the uniformly positive condition, followed by the negative to positive, positive to negative, and uniformly negative conditions.

Thus, research has not consistently supported either the gain-loss model or the reinforcement-affect model of interpersonal attraction. Nevertheless, research has suggested that reinforcing and punishing stimuli influence attraction by altering affective state.

Affective State

Research in interpersonal attraction generally supports both the reinforcement-affect and gain-loss models' assumption that environmental stimuli influence evaluative responses by changing affective state. Several studies have provided evidence that subjects' affective states are highly related to their attraction to a stranger (Gouax, 1971; Gouax and Lamberth, 1971; Griffitt, 1970; Wehmer and Izard, 1962).

Aronson and Linder (1965) reasoned that if affective state controls the relationship between overheard evaluative remarks about the subject and the subject's attraction for the evaluator, then the gain-loss ordering should occur only if the evaluator's negative evaluations produced a negative affective state in the subject. Accordingly, after they presented an evaluative sequence, an interviewer asked subjects in the negative-positive condition if they were bothered, embarrassed, or upset by the negative evaluation. Subjects who were upset by the negative evaluation liked the confederate at the end of the sequence more than subjects who were not upset. Similarly, in the positive-negative condition, those who were upset by the negative evaluation liked the confederate less than those who were not upset. These findings are consistent with the affective explanation of the gain-loss ordering because they suggest that negative affect is necessary for the greater liking in the negative-positive condition and for the greater disliking in the positive-negative condition. A fifth condition in the Aronson and Linder study was one in which the confederate's initial evaluation of the subject

was neutral and then became positive. The liking scores were almost identical to the positive-positive condition, again suggesting that negative affect is necessary for the gain effect.

Tognoli and Keisner (1972) in an extension of Aronson & Linder's 1965 study, collected a rudimentary measure of self-esteem by asking how subjects felt about themselves. Self-esteem scores showed the same group-to-group ordering as liking scores. That is, self-esteem was highest in the group in which liking was the highest, second highest in the group in which liking was the second highest, etc. Mettee, Taylor, and Friedman (1973, Study 1) found that subjects in the negative-positive condition greatly differed in anxiety at the end of the sequence. Subjects with low final anxiety liked the negative-positive evaluator whereas subjects with high final anxiety did not like the evaluator, suggesting that negative affect i.e. anxiety, is related to low liking within the negative-positive condition. In a second experiment, Mettee et al. (1973, Study 2) reported lower anxiety as well as higher liking for the negative-positive group than the positive-positive group at the end of the feedback sequence.

Application to the Classroom

Reward and punishment present in the classroom generate positive and negative affect along with low and high self-esteem (Weiner, 1979a, 1979b). The instructor may be perceived as responsible for and/or associated with the reinforcing and punishing events. Student evaluation of the instructor may thus be mediated by the positive and negative affect accomanying classroom reinforcement and punish-

ment. According to this analysis, lecturers associated with reinforcement will receive high ratings due to positive affect while those associated with punishment will receive low ratings due to negative affect. One primary classroom characteristic which may be viewed as rewarding or punishing is lecture quality. For example, students exposed to a poor lecture may experience negative feelings related to lack of understanding, perceived self-incompetence, anticipation of poor exam performance, boredom or frustration.

As noted previously, the reinforcement-affect model of interpersonal attraction maintains that attraction is a function of reward and punishment associated with the stimulus person. It proposes greatest attraction for persons associated with invariant reinforcing events and the least attraction for persons associated with invariant punishing events. Thus, the reinforcement-affect model predicts the highest student ratings for instructors who provide consistently good lectures and the lowest ratings for lecturers who provide consistently poor lecture quality.

The gain-loss model predicts student ratings that correspond to Leventhal et al.'s findings. A lecturer who produces in the student a gain in positive affect and self-esteem (poor lecture 1 quality, good lecture 2 quality) should provide more reward and thus be rated higher than an instructor who produces consistently positive affect and high self-esteem (good lecture 1 quality, good lecture 2 quality). Similarly, a lecturer who produces a loss in positive affect and self-esteem (good lecture 1 quality, poor lecture 2 quality) should provide more punishment and thus be rated lower than an instructor

who produces consistently negative affect and low self-esteem (poor lecture 1 quality, poor lecture 2 quality). Regardless of which ordering, reinforcement-affect or gain-loss, occurs in the classrroom, if measures of affect and self-esteem are related to student ratings, they should show the same ordering of lecture quality sequence as ratings.

SUMMARY

The relationship between lecture quality and student ratings may be viewed as an instance of the effect of reinforcing and punishing stimuli on interpersonal attraction. The reinforcement-affect model of interpersonal attraction proposes that stimuli associated with consistent reinforcement or punishment have the greatest effect on evaluative responses whereas the gain-loss model proposes that stimuli associated with gains or losses in reinforcement have the greatest effect. Research provides support for either model.

The ordering of student ratings found by Leventhal et al. is consistent with the gain-loss model. However, students were explicitly directed to evaluate the second portion of the lecture sequence (lecture 2). The major purpose of this study was to compare the effect on the gain-loss ordering of (a) instructions to evaluate the second lecture as in Leventhal et al. and (b) instructions which direct the subjects to consider both lecture 1 and lecture 2.

A secondary purpose of the study was to examine the relation of affective state and self-esteem to student ratings. Interpersonal attraction models and related research suggest that affective state mediates the relationship between reinforcement/punishment and inter-

personal evaluation. Measures of affect and self-esteem should show the same ordering of lecture quality sequence as student ratings if they are mediators. Another secondary purpose of the study was to compare results for overall liking with results for student ratings.

Thus, the study was conducted to answer the following research questions: (a) Do student rating instructions influence the gain-loss ordering of student ratings, (b) Do affect and self-esteem mediate between lecture quality sequence and ratings, and (c) Do instructions and lecture quality affect liking in the same way as ratings?

METHOD

Subjects

The subjects were 83 male and female students enrolled in two intersession psychology and two summer psychology courses at the University of Manitoba. Thus, the experiment was run in two replications. The students volunteered for the experiment without knowing its purpose. Within each semester, students selected two scheduled session times, two days apart. Each pair of sessions was then randomly assigned to experimental conditions. Three additional undergraduate psychology students volunteered to serve as judges to determine whether the research subjects were suspicious of the experimental manipulations.

Design

The design of the study was a 2 x 2 x 2 factorial. The variables manipulated across subjects, were lecture 1 quality (good, poor), lecture 2 quality (good, poor) and student rating instructions (consi-

der lecture 2, consider lecture 1 and lecture 2). The design may also be viewed as a 4 x 2 factorial: lecture quality sequence (poor-good, good-good, poor-poor, good-poor) by rating instructions. Dependent measures following lecture 2 were seven student rating scales of the instructor, liking for the instructor, affect, self-esteem and a quiz covering lecture 2. A quiz covering lecture 1 was administered after lecture 1.

Apparatus and Setting

The apparatus consisted of an Advent Model 1000A television monitor which projected a colour image on an approximately 7 x 5 ft. screen. Tapes were played on a Sony Model VO-2611 Videocasette recorder.

The setting was a standard size classroom seating 30. Students were seated on combination seat-desks facing the screen.

Materials

Instructions. Two sets of typed instructions prefaced a teacher rating form which was administered to the students after the second lecture. The instructions designed to direct students to rate lecture 1 included the following:

> Teacher rating forms are often given to students at the end of a course to assess teaching performance. Recall that you viewed two videotaped lectures representing the instructor's performance near the beginning and near the end of a university course. For this "end-of-course" teacher evaluation form we would like you to base your responses on the instructor's performance in the videotape of the

second lecture - the videotape you have just observed.

The instructions designed to direct students to consider both lecture 1 and lecture 2 included the following:

> Teacher rating forms are often given to students at the end of a course to assess teaching performance. Recall that you viewed two videotaped lectures representing the instructor's performance near the beginning and near the end of a university course. For this "end of course" teacher evaluation form we would like you to consider the instructor's performance in the videotapes of the first and second lectures.

Videotaped lectures. Eight 30-minute colour videotaped lectures by the same instructor were previously prepared. For this study four lectures were employed; two lectures (poor and good quality) on repression and two lectures (poor and good quality) on sex-role stereotypes. The quality of the lectures was manipulated by varying instructor expressiveness and lecture content. Using a procedure similar to Williams and Ware (1976) and Perry, Abrami, Leventhal and Check (1979), expressiveness was manipulated by varying the lecturer's enthusiasm, humour, voice inflection, eye contact and physical movement. Lecture content was manipulated by varying the clarity and number of teaching points contained in the lecture. The poor quality lectures consisted of both low expressiveness and low content while the good quality lectures consisted of both high expressiveness and high content.

Dependent Measures. Six measures were employed to investigate

the effects of the manipulations. First, a 30-item multiple-choice achievement test was given after lecture 1. The purpose of this quiz was to provide similar motivating influences to an actual classroom. Second, a similar quiz assessed student performance after lecture 2. The multiple-choice tests were similar to tests in undergraduate survey courses. Third, a 26-item teacher rating form measured student evaluations of the instructor after lecture 2. Item 1 of the rating form is an overall teaching ability item used by Sullivan and Skanes (1974). Items 2 - 26 were adapted from a teacher rating form constructed and factor analyzed by Hildebrand and Wilson (1970). The factors are as follows: analytic/synthetic approach - i.e. scholarship; organization/clarity; instructor-group interaction; instructor-individual-interaction; and dynamism/enthusiasm. All items were evaluated on a scale of 1 to 5, from unfavourable to favourable. Fourth, an overall liking measure at the end of the teacher rating form asked students to indicate "How much, in general, did you like the instructor as a person?" on a 5 point scale. This item is similar to the primary evaluation measure used in interpersonal attraction studies (e.g. Hewitt, 1972; Mettee, 1971; Tognoli and Keisner, 1972). Fifth, to assess the subject's affective state a "feeling scale" (Byrne and Clore, 1970; Byrne and Griffitt, 1969; Griffitt and Veitch, 1971) was administered following the liking measure. This measure consists of six of the evaluative scales of Osgood's Semantic Differential (Osgood, Suci and Tannenbaum, 1957). The scales are comfortable-uncomfortable, bad-good, high-low, sad-happy, pleasant-unpleasant and negative-positive. Students were asked to indicate

on a five point scale how closely the feelings correspond to their own. Lastly, after the affect scale, a short form of the Texas Social Behaviour Inventory (TSBI) provided a measure of self-esteem. The TBSI is a validated objective measure of self-esteem. The scale has been administered to more than 8,000 students and has been used in a variety of academic and non-academic settings (Helmreich and Stapp, 1974). The inventory consists of 16 statements which subjects rated on a five point scale as characteristic or not characteristic of themselves.

Post-lecture questionnaire. The first item of this questionnaire served as a check on the effectiveness of the instructional manipulation. It asked the students on which lecture(s) they based their evaluation responses. The next four items, which constituted a "suspiciousness questionnaire", assessed whether subjects guessed the purpose of the experiment and whether their suspicions about the experiment influenced their responses.

Procedure

The study consisted of eight first sessions and eight second sessions. At the beginning of the first session, all students were told that the experiment was concerned with studying instruction under laboratory conditions and that they were to view a lecture delivered by an instructor who has volunteered to participate in the experiment and write a short quiz on the lecture. They were asked to return two days later to view a second lecture, write a quiz, evaluate the instructor, and complete a few student information forms. Students were asked to consider the first lecture as being delivered

near the beginning of a university course and the second lecture as being delivered near the end of a university course. They then viewed either the good or poor quality lecture on repression and completed the quiz. All groups were provided with writing materials and permitted to take notes during the lectures. At the beginning of the second session, the general instructions and tasks required were repeated. Students then viewed either a good or poor quality lecture on sex-role stereotypes and completed the second quiz. After the quiz, half the groups were directed to evaluate the instructor's performance during lecture 2 while the other groups were asked to consider both lecture 1 and 2 in their evaluations of the instructor. The global liking measure was included at the end of the teacher rating form. All students then completed the affect and self-esteem measures and the post-lecture questionnaire. Students received a written and verbal debriefing which explained the manipulations and the purpose of the experiment either in their classrooms or by mail at the conclusion of all eight sessions. Students were also phoned to ensure that they had received and understood the debriefing.

RESULTS

Subject Attrition and Suspiciousness

Three judges rated the suspiciousness questionnaires, using a three-point scale where 1 meant the student was unsuspicious, 2 meant the judge was uncertain, and 3 meant the student was suspicious. Only one student received a total of over five points and was excluded from the analysis. Two other students failed to attend the second session. The following analyses apply to the remaining 80 students.

Student Variables

Students in the experiment differed in several ways. Students were either North American or foreign, and came from either intersession or summer session courses. If student differences were not represented equally among the eight treatment conditions, differences in treatment conditions could have been due to student differences rather than the experimental manipulations.

Session. Students were classified according to which session they attended (intersession or summer). A chi-square test of association on a 2 x 8 contingency table showed that students from the two sessions were not represented equally among the eight treatment conditions, $\chi^2(7) = 33.44$, $p < .05$. However, one-way ANOVA's showed no significant differences on any of the dependent measures between intersession and summer students ($p > .05$). Hence, it is concluded that the differences among treatment conditions on dependent measures were not due to session.

Language and origin. Students were classified according to first language spoken and continent of origin. Forty-six students spoke English as a first language (category English), 24 first spoke Chinese or Malaysian (category Asian), and 10 spoke European, African, or Cree (category Other).

A chi-square test of association on each of two 3 x 8 contingency tables indicated that students in the three categories of language and three categories of country did not differ significantly among the eight lecture conditions, $\chi^2(14) = 12.35$, $p > .10$; $\chi^2(14) = 15.87$, $p > .10$ respectively. It is therefore concluded that dif-

ferences among treatment conditions on dependent measures were not the result of student differences in first language spoken or continent of origin.

Manipulation Checks

If the manipulation of lecture 1 quality was successful, lecture 1 would be expected to influence student achievement on quiz 1. A Lecture 1 x Lecture 2 x Instructions ANOVA indicated that lecture 1 quality had a significant and strong effect on lecture 1 quiz scores, $F(1,72) = 28.17$, $p < .0001$, $\omega^2 = .25$. No other main effect or interaction was significant, which was anticipated. Since both lecture 2 and the teacher rating instructions occurred after lecture 1, they should not have affected lecture 1 quiz scores. Lecture 2 quiz scores were also analyzed with a Lecture 1 x Lecture 2 x Instructions ANOVA. Lecture 2 quality significantly and strongly affected lecture 2 quiz scores, $F(1,72) = 48.04$, $p = < .0001$, $\omega^2 = .37$. No other main effect or interaction was significant. Apparently, lecture 1 and lecture 2 quality greatly affected student achievement.

In the post-lecture questionnaire, students were asked to indicate on a scale of 1 to 5 which lecture(s) they based their evaluation on (1 = lecture 1, 2 = mostly lecture 1, slightly lecture 2, 3 = lecture 1 and lecture 2, 4 = mostly lecture 2, slightly lecture 1, 5 = lecture 2). If the instructional manipulation was effective, the instructions variable should show a large effect on this lecture-considered measure. A Lecture 1 x Lecture 2 x Instructions ANOVA on lecture-considered data indicated that instructions significantly and strongly affected the lecture students considered when evaluating

the instructor, $F(1,72) = 85.91$, $p < .0001$, $\omega^2 = .51$. No other main effect or interaction was significant.

Primacy-Recency Effects

For the reader's convenience, Table 1 presents the cell means of ratings, liking, affect, and self-esteem for all treatment conditions, not only primacy and recency effects. Primacy-recency effects

---

Insert Table 1 about here

---

were assessed for each of the lecture 2 dependent measures by performing, for each measure, a Lecture 1 x Lecture 2 x Instructions ANOVA. For any measure, a lecture 1 main effect was a primacy effect and a lecture 2 main effect was a recency effect. The results of the ANOVA's, and ω^2 values, are presented in Table 2.

Insert Table 2 about here

Student ratings. The major research question was whether instructions influence the lecture quality sequence ordering of student ratings. Summarizing the significant effects for ratings in Table 1, both lecture 1 and lecture 2 significantly interacted with instructions for all seven rating measures. This indicated that the effect of lecture 1 and of lecture 2 on ratings depended on whether instruc-

tions were to consider lecture 2 or lecture 1 and 2. There was a significant main effect of lecture 2 for all rating measures. Instructions had a significant main effect for six of the seven rating measures.

To determine the effects of lecture 1 and lecture 2 on ratings under each level of instructions, Lecture 1 x Lecture 2 ANOVA's were conducted separately for lecture 2 instructions and lecture 1 and 2 instructions. Primacy effects are indicated by lecture 1 effects and recency effects are indicated by lecture 2 effects. Table 3 presents the results of the ANOVA's and ω^2 values. Means for lecture

---

Insert Table 3 about here

---

1 and lecture 2 for each instruction level are shown in Table 4.

Insert Table 4 about here

As shown in Table 3, when students were instructed to consider only lecture 2 in their ratings, primacy effects, small to moderate in strength, were present for six of the seven rating measures. Very strong recency effects were present for all rating measures. Inspection of the means in Table 4 for the poor and good lecture 1 shows the primacy effect to be <u>negative</u>; that is, poor lecture 1 subjects

gave significantly higher ratings than good lecture 1 subjects. The recency effects were positive; that is, good lecture 2 subjects gave significantly higher ratings than poor lecture 2 subjects.

As seen in Table 4, when instructions were to consider lecture 1 and 2, the primacy effects, small to moderate in strength, were positive. That is, ratings were higher for good lecture 1 subjects than poor lecture 1 subjects. The recency effects also were positive; that is, good lecture 2 subjects gave significantly higher ratings than poor lecture 2 subjects. Table 3 illustrates that the positive primacy effect reached statistical significance for three of the rating measures. The positive recency effect was significant for all the rating measures.

The Lecture 1 x Instructions interaction thus resulted from negative primacy effects under lecture 2 instructions and less robust positive primacy effects under lecture 1 and 2 instruction. The Lecture 2 x Instruction interaction resulted from larger positive recency effects under lecture 2 instructions than lecture 1 and 2 instructions. As can be seen in Table 3 and Table 4, primacy effects were consistently less strong as assessed by ω^2, and smaller as assessed by the size of the mean difference between good lecture 1 and poor lecture 1, than recency effects.

In summary, under lecture 2 instructions, there were: (a) consistent negative primacy effects, significant and small to moderate in strength and size for six out of seven ratings, and (b) consistant positive recency effects, significant and very large in strength and size for all ratings. Under lecture 1 and 2 instructions, effects

were (a) consistent positive primacy effects, significant and small to moderate in strength and size for three ratings, and (b) consistent positive recency effects, significant and large in strength and size for all ratings.

Liking. A secondary research question was whether lecture quality and instructions have similar effects on ratings and liking. As seen in Table 2, there was a significant Lecture 1 x Instruction interaction which indicated that the effect of lecture 1 on liking depended on instructions. The effect of lecture 1 at each level of instructions is shown in Table 3. The negative primacy effect under lecture 2 instructions was significant and moderate in strength. The positive primacy effect under lecture 1 and 2 instructions was not significant.

Table 2 shows a significant main effect of lecture 2 in liking. Unlike the effect of lecture 2 on ratings, recency effects on liking were not significantly larger under lecture 2 instructions than lecture 1 and 2 instructions; that is, there was no Lecture 2 X Instructions interaction on liking. Recency effects were positive and large in strength.

To summarize, liking, in general, showed the same effects as ratings. Under lecture 2 instruction, there was a negative primacy effect, significant and moderate in strength and size, and a positive recency effect, significant and large in strength and size. Under lecture 1 and 2 instructions, there was an indication of a positive primacy effect which was not significant and a positive recency effect, significant and large in strength and size. Unlike ratings, the size of the recency effect did not significantly differ depending

on instructions.

Affect. Another secondary research question was whether affect and self-esteem mediate between lecture quality sequence and student ratings, and hence show the same lecture quality sequence ordering as ratings. Inspection of the means of affect in Table 4 illustrates that affect showed the same primacy-recency pattern as ratings. However, only some of the effects were significant. Table 2 shows a significant main effect of lecture 2 on affect. In addition, a significant Lecture 1 x Lecture 2 interaction indicated that the effect of lecture 1 on affect depended on whether lecture 2 was good or poor. Probes of the interaction revealed two significant comparisons. First, when lecture 2 was poor, affect was higher for the good lecture 1 than the poor lecture 1, (poor-poor $<$ good-poor, q (2,72) = 4.06, $p < .01$). Second, when lecture 1 was poor, affect was higher for the good lecture 2 than the poor lecture 2 (poor-good $>$ poor-poor, $q(2,72) = 5.03$, $p < .01$). These effects have implications for the gain-loss ordering which are discussed in the next section.

To summarize, there was a positive primacy effect on affect when lecture 2 was poor but not when lecture 2 was good and a larger positive recency effect when lecture 1 was poor than when it was good.

Self-esteem. Self-esteem showed only a main effect of lecture 1, with self-esteem scores significantly higher when lecture 1 was poor than when lecture 1 was good. This amounted to an overall negative primacy effect.

Gain-Loss Effects

Ratings and liking. The negative primacy effect and the larger

positive recency effect on student ratings and liking under lecture 2 instructions produce for each measure a gain-loss ordering (poor-good, good-good, poor-poor, good-poor). The positive primacy effect and larger positive recency effect on student ratings and liking under lecture 1 and 2 instructions produce for each measure a reinforcement-affect ordering (good-good, poor-good, good-poor, poor-poor). These orderings are shown in Table 1.

From a gain-loss perspective, the gain-loss ordering under lecture 2 instructions includes the gain effect poor-good > good-good) and the loss effect (poor-poor > good-poor), both of which may be viewed as negative primacy effects. Both gain and loss effects on ratings and liking occurred under lecture 2 conditions. Furthermore, although inspection of means in Table 4 indicates that gain effects were consistently larger than loss effects, gain and loss effects did not differ significantly. If they had, Lecture 1 x Lecture 2 interactions would have been evidenced. This is so because what distinguishes gain and loss effects, which are both the result of negative primacy effects (lecture 1 effects), is lecture 2 quality. A negative primacy effect when lecture 2 quality is good is a gain effect. A negative primacy effect when lecture 2 quality is poor is a loss effect. As seen in Table 3, a Lecture 1 x Lecture 2 interaction occurred for only one rating measure.

A similar analysis may be made for the reinforcement-affect ordering under lecture 1 and 2 instructions. The reinforcement affect ordering includes the invariant reinforcement effect (good-good > poor-good) and the invariant punishment effect (good-poor>

poor-poor), both of which may be viewed as positive primacy effects. Positive primacy effects consistently occurred under lecture 1 and 2 instructions and were significant for three rating measures. Effects of both invariantly good and invariantly poor lectures occurred and did not differ significantly in magnitude. If they had, significant Lecture 1 x Lecture 2 interactions would have been evidenced. This is so because the effects of invariantly poor and invariantly good lectures, which are both the result of positive primacy effects (lecture 1 effects), are distinguished by lecture 2 quality. A positive primacy effect when lecture 2 quality is good is an invariantly good lecture effect. A positive primacy effect when lecture 2 quality is poor is an invariantly poor lecture effect.

In summary, both gain and loss effects were consistently present under lecture 2 instructions. The effects of both invariantly good and invariantly poor lectures were consistently present under lecture 1 and 2 instructions.

Affect. If affect resulting from lecture quality sequence influences ratings and liking, the ordering of affect, ratings and liking across lecture quality sequence should be similar. Thus, under each level of instructions, the ordering of lecture sequence for affect was compared to the ordering for ratings and liking.

As seen in Table 1, affect shows a partial correspondence to ratings and liking under both lecture 2 and lecture 1 and 2 instructions. Under lecture 1 and 2 instructions, affect's order of the good-good and poor-poor sequences was respectively, rank 1 and 4, the same as for ratings and liking. But affect's order of the poor-

good and good-poor sequences was rank 3 and 2, which was the reverse of the ranking found for liking and ratings. Under lecture 2 instructions, affect's rank order followed that for liking and ratings except for poor-poor and good-poor which, for affect, ranked 4 and 3, and was the reverse of the ranking found for liking and ratings.

DISCUSSION

The major purpose of this study was to investigate the effect on the gain-loss ordering of student ratings of instructions to students to consider lecture 2 as compared with instructions to consider both lecture 1 and lecture 2. Secondary purposes were to determine whether affective state and self-esteem of students mediate between lecture quality sequence and student ratings of the instructor, and to determine if lecture quality affects liking in the same way as it affects ratings.

As stated in the introduction, lecture quality is a classroom characteristic which may be viewed as reinforcing or punishing to the student. Student ratings and liking of the instructor may be viewed as measures of interpersonal attraction.

Liking

The liking measure showed the same primacy and recency effects as ratings with one exception. The recency effect was significantly larger under lecture 2 instructions than lecture 1 and 2 instructions for ratings but not for liking. The lecture quality sequence ordering under each instruction level for liking corresponded

to the ordering for ratings. Thus, liking appears to be affected by one classroom reinforcer, namely lecture quality, in the same way that ratings are affected. One implication is that instructors who are good teachers are not only recognized for their teaching ability, they are also attributed with superior personal characteristics.

Instructions

Different student rating and liking orderings were predicted by the reinforcement-affect and the gain-loss models of interpersonal attraction. When students were instructed to consider both lecture 1 and 2 in their evaluations, positive primacy effects and larger positive recency effects on ratings and liking occurred, resulting in a reinforcement-affect ordering (good-good, poor-good, good-poor and poor-poor lecture quality).

When students were instructed to consider lecture 2 only when evaluating the teacher, negative primacy effects and larger positive recency effects on ratings and liking occurred, resulting in a gain-loss ordering (poor-good, good-good, poor-poor and good-poor lecture quality). Gain effects (poor-good $>$ good-good) were consistently larger than loss effects (poor-poor $>$ good-poor), although the difference was not statistically significant. This finding is consistent with the interpersonal attraction literature in which gain effects have traditionally been stronger than loss effects. Even in the original Aronson and Linder (1965) gain-loss study, the loss effect was weaker than the gain effect.

Replacement versus addition of information. The effects of the two kinds of instructions on student ratings and liking may be related to whether later information is perceived as replacing or adding to earlier information. For example, the information that you are not intelligent replaces the information that you are intelligent. The information that you are charming adds to the information that you are intelligent. New information replaces prior information when it is logically contradictory and cannot co-exist with prior information. Mettee and Aronson (1974) point out that the gain-loss effect has been demonstrated only when recent information is perceived as replacing prior information. Mettee, Taylor and Friedman (1973) propose that the gain effect occurs when new information, in addition to establishing reinforcement and its accompanying affective state, also reduces the negative affect produced by prior information. The loss effect occurs in a similar but opposite fashion. In order to reduce the effect of prior information, new information must be perceived as replacing prior information. When later information is perceived as an addition to early information, affect from the early and later information add together, resulting in the good-good condition being the most attractive and the poor-poor condition the least attractive. This produces the reinforcement-affect ordering.

Applied to the present data, Mettee et al.'s hypothesis suggests that when students were instructed to consider lecture 1 and 2, they may have perceived later lecture quality to be an addition to earlier lecture quality. As proposed earlier, students

experience positive or negative effect in response to lecture quality. In the poor to good sequence, negative affect is elicited by the poor initial lecture and positive affect is elicited by the good later lecture. The positive and negative affect tend to counteract each other and produce less liking and lower ratings than a good-good sequence in which the positive affect from both good lectures add together. Similarly, good-poor teaching produces more liking and higher ratings than poor-poor teaching. In contrast, when students are told to attend only to lecture 2, they are left with positive affect from lecture 2 plus additional positive affect resulting from the removal of negative affect associated with lecture 1. The student exposed to two good lectures experiences only the positive affect produced by the second lecture. The result is the gain effect in which liking and ratings for poor-good exceed good-good. A similar process occurs for the loss effect.

A problem with this explanation of how instructions can produce either gain-loss or reinforcement-affect orderings is that the explanation depends on affect as a mediator between lecture quality and ratings. As will be pointed out later, the present data provided only moderate evidence of mediation by affect.

Implications for the classroom. The implications of these findings for the classroom are that when students are implicitly or explicitly told to consider teaching performance occurring later in the course, instructors who improve will be rated more highly and liked more than instructors who are consistently good.

Instructors whose teaching performance deteriorates will be given lower ratings and liked less than those whose performance is consistently poor. When students are told to consider lecture quality throughout a course, instructors who are consistently good will receive higher ratings and will be liked more than those whose performance is initially poor and becomes good later. Instructors who demonstrate consistently poor lecture quality will be rated lower and liked less that those whose performance is initially good and becomes poor later. Although we may speculate that students generally are implicitly led to consider teaching performance throughout a course, we have no evidence of this. It would thus be of interest to determine to what extent each of the instruction conditions are found in the classroom and relatedly which portions of a course students consider when evaluating teaching performance.

Implications for interpersonal attraction literature. An issue in the interpersonal attraction literature is whether research has best supported reinforcement-affect theory or gain-loss theory. Previous inconsistencies in the literature may be partially due to differences in the stated or implied emphasis on recent information or previous and recent information. This study suggests that when later information is emphasized, later information has a large positive effect and early information has a smaller negative effect, resulting in the gain-loss ordering. When both early and later information are emphasized, later information has a large positive effect and early information has a smaller positive effect, resulting in the reinforcement-affect ordering.

Affect

Both the gain-loss and reinforcement-affect models propose that attraction is influenced by reinforcement and punishment through affective state. Thus, when affect is great, attraction should be great, and when affect is low, attraction should be low. Under both instruction conditions, the order of lecture quality sequences for affect showed a partial correspondence to the order of sequences for liking and ratings. Thus, there is a suggestion that affect plays a role in the relationship between lecture quality and student ratings.

The relationship among affect, lecture quality and ratings may be weak due to one of two reasons: First, affect may be a mediator but this study may not have been a fair test of it. Table 1 indicates that the affect measure employed here was not very sensitive to the experimental manipulations. Second, affect may not be a mediator and the present findings are due to chance.

Self-esteem

The only effect involving self-esteem was lecture 1 quality. Self-esteem was higher when lecture 1 quality was poor than when it was good! Perhaps rather than thinking less of themselves because of how little they understood or learned from the instructor, students compared themselves to their first impression of the inept instructor and found the differences flattering.

There was no evidence in this study that self-esteem was related to student ratings or liking. Tognoli and Keisner (1972)

found that self-esteem showed the same ordering of sequences as liking. However, their experiment differed in at least two relevant ways from this one. First, in their study, reinforcement and punishment consisted of favourable and unfavourable evaluations about the subject. This may affect self-esteem more than lecture quality. Second, their self-esteem measure referred specifically to the experimental situation (e.g. "How did you feel about yourself as a result of what the other person said about you both at the beginning and at the end of the experiment?"). In contrast, the self-esteem inventory used in this study (TSBI) may be viewed as a trait measure which would be affected little by isolated environmental events.

An Alternative Explanation of the Gain-Loss Ordering

Research in areas other than gain-loss has suggested the greater potency of change over consistency in the presentation of reinforcement and punishment. Most explanations employ the concept of contrast, for example, behavioural contrast in operant conditioning, positive and negative contrast in instrumental learning (see footnote 1), and adaptation-level contrast (Helson, 1964).

Perhaps of particular relevance to the present study is judgemental contrast as proposed by Berkowitz (1960a and 1960b). Berkowitz found that a partner's friendly or hostile act had its greatest effect on a subject's own final attitude of friendliness or hostility toward the partner when the act was inconsistent with the partner's earlier behaviour. If the partner were hostile initially, but in a second communication were friendly, the sub-

ject's friendliness towards him increased more than if the partner had been consistently friendly. Similarly, if the partner were initially friendly, but were hostile in a second communication, the subject's hostility towards him increased more than if the partner had been consistently hostile. Berkowitz postulated that the reinforcement or punishment initially received served as a comparison standard in judging the degree of later reinforcement or punishment, resulting in the subject's own friendliness or hositility. For example, recent reinforcement was perceived as greater when it was compared to past punishment than to past similar reinforcement.

Extending this analysis to the classroom situation, ratings and liking may have been higher for the poor-good sequence than the good-good sequence and lower for the good-poor sequence than the poor-poor sequence under lecture 2 instructions because the reinforcement value of the second lecture, which was the lecture being evaluated, was compared to that of the first lecture. In the poor-good situation, the second lecture was compared favourably to that of the first lecture. In the good-good situation, the good second lecture equalled the first lecture standard. Similarly, in the good-poor situation, the second lecture was compared unfavourably to that of the first lecture. In the poor-poor situation, the poor second lecture equalled the standard.

Aronson and Linder briefly considered contrast as an alternative explanation for their findings but dismissed it. They reasoned that if contrast were operating it should be evident in a neutral-

positive condition. Liking in a neutral-positive condition was lower than in the negative-positive condition and higher than in the positive-positive condition. However, liking in the neutral-positive condition was closer to liking in the positive-positive condition than the negative-positive condition, suggesting to them contrast effects were not operating.

Nevertheless, a contrast explanation for their gain-loss ordering can not be unequivocally rejected. First, liking in the neutral-positive condition was lower than in the negative-positive sequence and higher than in the positive-positive sequence with neither difference being significant. This is the rank order predicted by contrast. Second, Aronson and Linder's neutral condition consisted of negative and positive feedback about the subject. Subjects may have been more influenced by the positive than the negative statements, making the neutral-positive condition more similar to the positive-positive condition than the negative-positive condition. This would result in liking scores which were closer to those in the positive-positive than negative-negative condition, as found. Third, a neutral-negative condition, which was not studied, may have shown more support for the contrast interpretation.

Mettee et al.'s replacement-addition hypothesis together with a contrast hypothesis can account for how instructions produce either the gain-loss or reinforcement-affect ordering. When students are instructed to consider lecture 2, they determine the reinforcement value of the lecture by using the first lecture

as a comparison standard, resulting in the gain-loss or "contrast" ordering. When students are instructed to consider both lecture 1 and 2, the reinforcement value is based on both lectures, with the reinforcement value of lecture 1 added to that of lecture 2, resulting in the reinforcement-affect ordering.

Summary

The three research questions were: (a) Do student rating instructions influence the gain-loss ordering of student ratings, (b) Do affect and self-esteem mediate between lecture quality sequence and ratings, and (c) Do instructions and lecture quality affect liking in the same way as ratings?

Briefly, it was found that:

1. Student rating instructions influenced the lecture quality sequence ordering of student ratings. When instructions were to consider lecture 2, ratings showed a gain-loss ordering; that is, ratings were highest for the poor-good lecture quality sequence, followed by the good-good, poor-poor, and good-poor sequences. When instructions were to consider lecture 1 and 2, ratings showed a reinforcement-affect ordering; that is, ratings were highest for the good-good lecture quality sequence, followed by the poor-good, good-poor and poor-poor sequences.

2. Instructions and lecture quality affected liking the same as ratings; thus, liking and ratings showed the same ordering.

3. There was a suggestion that affect mediates between lecture quality and student ratings and between lecture quality and liking, but the evidence was inconclusive.

Research Implications

The effect of instructions on the lecture quality sequence ordering of ratings and liking may relate to whether lecture 2 quality was perceived as replacing or adding to lecture 1 quality. This explanation holds that, in the poor to good sequence under lecture 1 and 2 instructions, negative affect from lecture 1 and positive affect from lecture 2 tended to counteract each other and produce less liking and lower ratings than a good-good sequence. Under lecture 2 instructions, students were left with positive affect from lecture 2 plus additional positive affect resulting from the removal of negative affect associated with lecture 1. This resulted in greater liking and ratings for the poor-good sequence than the good-good sequence. The replacement-addition explanation depends on affect as a mediator between lecture quality and ratings and between lecture quality and liking. This study suggested that affect was a mediator but the evidence was inconclusive. Affect was previously proposed to be influenced by lecture quality partially because students anticipate poor or good academic achievement. The role of affect could be tested by experimentally manipulating the importance of student quiz performance. Affect should be stronger when achievement is made more important and weaker when achievement is made less important. Thus, in the replacement situation (lecture 2 instructions), when affect is strong, gain-loss ratings should occur. But when affect is weak, the gain-loss ordering should be less evident. In the

addition situation (lecture 1 and 2 instructions), the reinforcement-affect ordering should be more evident when affect is strong than when it is weak.

Studies with pre-arranged replacement and addition situations can be conducted to test further the replacement-addition hypothesis. For example, Leventhal et al. suggested manipulating the similarity of lecture 1 and 2. In the replacement condition, the same topic and same instructor for both lecture would be used. In the addition condition, lecture 2 would be a different topic delivered by a different instructor. If the replacement-addition explanation for the results of this study is correct, a gain-loss ordering should occur in the high similar condition and a reinforcement-affect ordering should occur in the low similar condition.

An alternative explanation for the gain-loss ordering and the results of this study involves contrast effects. This explanation holds that, under lecture 2 instructions, the reinforcement value of the second lecture is determined by comparing it with the first lecture, resulting in the gain-loss ordering. Under lecture 1 and 2 instructions, the reinforcement value is based on both lectures, resulting in the reinforcement-affect ordering. The contrast explanation can be tested by including a neutral (neither good nor poor) lecture 1. Under lecture 2 instructions, a sequence comprised of a neutral lecture 1 and a good lecture 2 should produce less contrast than a poor-good sequence and more contrast than a good-good sequence. Thus, ratings for the neutral-good

lecture sequence should be lower than the poor-good sequence and higher than the good-good sequence. Similarly, ratings for a neutral-poor lecture sequence should be higher than the good-poor sequence and lower than the poor-poor sequence. Information regarding the processes involved in rating instruction may also be extracted by directly asking students in each instruction condition if they based their ratings on a comparison of lecture 1 and lecture 2.

Footnotes

1. It should be noted that this cognitive-reinforcement model does not represent all learning theory positions. For example, the phenomenon of contrast and behavioural contrast in instrumental and operant conditioning supports the notion that subjects shifted in reinforcement or punishment behave differently than subjects with constant reinforcement or punishment. In instrumental conditioning, "positive contrast" refers to the superior performance of animals in a runway who are shifted from small to large magnitude of reward, as compared to a control group receiving consistently large reward. Similarly, "negative contrast" refers to poorer performance when shifted from large to small magnitude of reward, as compared to consistently small reward (e.g. see Dunham's 1968 review). In operant conditioning, "behavioural contrast" occurs in an alternating two component, multiple schedule of reinforcement when a change in response rate appears in one schedule of reinforcement in a direction away from the response rate generated during the earlier schedule of reinforcement (e.g. Reynolds, 1961; Reynolds and Catania, 1961).

2. The gain-loss finding is supported by other research which suggests the greater potency of change versus consistency in the presentation of positive and negative stimuli. For example, several studies have found that attraction responses to agreement from a stranger which followed a series of disagreements were much more positive than agreement following a series of prior agreements (Gerard & Greenbaum, 1962; Stapert & Clore, 1969; and Worchel & Schuster, 1966). Other research is discussed in the Discussion section.

Reference Notes

1. Leventhal, L., Turcotte, S., Abrami, P., and Perry, R. Primacy/recency effects in student ratings of instruction (a reinterpretation of gain-loss effects). Manuscript submitted for publication, 1982.

2. Ramagli, H. and Greenwood, G. Another assessment of the Dr. Fox effect: Paired comparisons of expressiveness and content. Paper presented at the American Educational Research Association, Boston, April, 1980.

References

Aronson, E. Some antecedents of interpersonal attraction. In W. J. Arnold & D. Levine (Eds.), Nebraska Symposium on Motivation, (Vol. 17). Lincoln: University of Nebraska Press, 1969.

Aronson, E. & Linder, D. Gain and loss of esteem as determinants of interpersonal attractiveness. Journal of Experimental Social Psychology, 1965, 1, 156 - 171.

Berkowitz, L. Some factors affecting the reduction of overt hostility. Journal of Abnormal and Social Psychology, 1960a, 60, 14 - 21.

Berkowitz, L. Repeated frustration and expectations in hostility arousal. Journal of Abnormal and Social Psychology, 1960b, 60, 422 - 429.

Bersheid, E., Brothen, T. & Graziano, W. Gain-loss theory and the "law of infidelity": Mr. Doting versus the admiring stranger. Journal of Personality and Social Psychology, 1976, 33, 709 - 718.

Bramel, D. Interpersonal attraction, hostility and perception. In J. Mills (Ed.), Experimental social psychology. London: Macmillan, 1969.

Byrne, D. The attraction paradigm. New York: Academic Press, Inc., 1971.

Byrne, D. & Clore, G. A reinforcement model of evaluative responses Personality, 1970, 1, 103 - 128.

Byrne, D. & Griffitt, W. Similarity and awareness of similarity of personality characteristics. Journal of Experimental Research in Personality, 1969, 3, 179 - 186.

Byrne, D. Lamberth, J., Palmer, J. & London, O. Sequential effects as a function of explicit and implicit interpolated attraction responses. Journal of Personality and Social Psychology, 1969, 13, 70 - 78.

Byrne, D. & London, O. Primacy-recency and the sequential presentation of attitudinal stimuli. Psychonomic Science, 1966, 6, 192 - 194.

Clair, M. & Snyder, C. Effects of instructor-delivered sequential evaluative feedback upon students' subsequent classroom-related performance and instructor ratings. Journal of Educational Psychology, 1979, 71, 50 - 57.

Clore, G. & Byrne, D. A reinforcement-affect model of attraction.

In T. L. Huston (Ed.) Foundations of interpersonal attraction. New York: Academic press, Inc., 1974.

Clore, G., Wiggins, N. & Itkin, S. Gain and loss in attraction Attributes from nonverbal behaviour. Journal of Personality and Social Psychology, 1975, 31, 706 - 712.

Coutts, L., Schneider, F. and Montgomery, S. An investigation of the arousal model of interpersonal intimacy. Journal of Experimental Social Psychology, 1980, 16, 545 - 561.

Dixon, W. J. (Ed.). BMDP biomedical computer programs. Berkeley: University of California Press, 1979.

Gerard, H. & Greenbaum, C. Attitudes toward an agent of uncertainty reduction. Journal of Personality, 1962, 30, 485 - 495.

Gouax, C. Induced affective states and interpersonal attraction. Journal of Personality and Social Psychology, 1971, 20, 37 - 43.

Gouax, C. & Lamberth, J. Interpersonal attraction as a function of Izard's FIRS evaluations and affective states. Personality, 1971, 2, 289 - 297.

Griffitt, W. Context effects in response to affective stimuli. Personality, 1971, 2, 23 - 33.

Griffitt, W. & Veitch, R. Hot and crowded: Influences of population density and temperature on interpersonal affective behaviour. Journal of Personality and Social Psychology, 1971, 17, 92 - 98.

Helmreich, R & Stapp J. Short forms of the Texas Social Behaviour Inventory. Bulletin of the Psychonomic Society, 1974, 4, 473 - 475.

Helson, H. Adaptation-level theory. New York: Harper & Row, 1964.

Hewitt, J. Liking and the proportion of favourable evaluations. Journal of Personality and Social Psychology, 1972, 22, 231 - 235.

Hildebrand, M., & Wilson, R. Effective university teaching and its evaluation. Washington, D.C.: U.S. Department of Health, Education & Welfare, Office of Education, 1970.

Leach, C. The importance of instructions in assessing sequential effects in impression formation. British Journal of Social and Clinical Psychology, 1974, 13, 151 - 156.

Leventhal, L., Perry, R. & Abrami, P. Effects of lecture quality

and student perception of lecturer's experience on teacher ratings and student achievement. Journal of Educational Psychology, 1977, 69, 360 - 374.

Luchins, A. Experimental attempts to minimize the impact of first impressions. In C.I. Hovland (Ed.), The order of presentation in persuasion. New Haven: Yale University Press, 1957.

Mettee, D. Changes in liking as a function of the magnitude and affect of sequential evaluations. Journal of Experimental Social Psychology, 1971, 7, 157 - 172.

Mettee, D. & Aronson, E. Affective reactions to appraisal from others. In T.L. Huston (Ed.), Foundations of interpersonal attraction. New York: Academic Press Inc., 1974.

Mettee, D., Taylor, S. & Friedman, H. Affect conversion and the gain-loss liking effect. Sociometry, 1973, 36, 505 - 519.

Osgood, C., Suci, G. & Tannenbaum, P. The measurement of meaning. Urbana: University of Illinois Press, 1957.

Perry, R., Abrami, P., Leventhal, L. & Check, J. Instructor reputation: An expectancy relationship involving student ratings and achievement. Journal of Educational Psychology, 1979, 71, 776 - 787.

Reynolds, G. Behavioural contrast. Journal of the Experimental Analysis of Behaviour, 1961a, 4, 57 - 76.

Reynolds, G. and Catania, A. Behavioural conrast with fixed-interval and low-rate reinforcement. Journal of the Experimental Analysis of Behaviour, 1961, 4, 387-391.

Rosencrantz, P & Crockett, W. Some factors influencing the assimilation of disparate information in impression formation. Journalof Personality and Social Psychology, 1965, 2, 397 - 402.

Sigall, H. & Aronson, E. Opinion change and the gain-loss model interpersonal attraction. Journal of Experimental Social Psychology, 1967, 3, 178 - 189.

Stapert, J. & Clore, G. Attraction and disagreement produced arousal. Journal of Personality and Social Psychology, 1969, 13, 64 - 69.

Sullivan, A. & Skanes, G. Validity of student evaluations of teaching and the characteristics of successful instructors. Journal of Educational Psychology, 1974, 66, 584 - 590.

Taylor, D., Altman, I. & Sorrentino, R. Interpersonal exchange as a function of rewards and costs and situational factors:

Expectancy confirmation - disconfirmation. Journal of Experimental Social Psychology, 1969, 5, 324 - 339.

Tognoli, J. & Keisner, R. Gain and loss of esteem as determinants of interpersonal attraction: A replication and extension. Journal of Personality and Social Psychology, 1972, 27, 372 - 382.

Wehmer, G. & Izard, C. The effect of self-esteem and induced affect on interpersonal perception and intellective functioning. ONR Technical Report No. 10, Vanderbilt University, 1962.

Weiner, B. A theory of motivation for some classroom experiences. Journal of Educational Psychology. 1979a 71, 3 - 25.

Weiner, B. The cognition-emotion process in achievement-related contexts. Journal of Personality and Social Psychology, 1979b, 37, 1211 - 1220.

Williams, R. & Ware, J. Validity of student ratings of instruction under different incentive conditions: A further study of the Dr. Fox effect. Journal of Educational Psychology, 1976, 68, 48 - 51.

Winer, B. Statistical principles in experimental design. New York: McGraw- Hill Book Co., 1962.

Worchel, P. & Schuster, S. Attraction as a function of the drive state. Journal of Experimental Research in Personality, 1966, 1, 277 - 281.

Table 1

Lecture Quality Sequence
Cell Means and Standard Deviations for
Seven Rating Measures, Liking, Affect and Self-Esteem

		Lecture 2 Instructions				Lecture 1 & 2 Instructions			
		Poor-Good	Good-Good	Poor-Poor	Good-Poor	Good-Good	Poor-Good	Good-Poor	Poor-Poor
Overall	M	3.91	2.80	1.58	1.22	2.60	2.33	1.80	1.22
	SD	.83	.79	.51	.44	.70	.50	.63	.44
H+W 1	M	3.47	3.20	2.23	1.82	2.76	2.78	2.30	1.84
	SD	.71	.57	.67	.29	.46	.51	.58	.51
H+W 2	M	3.84	3.18	2.18	1.78	3.32	2.67	2.40	1.78
	SD	.36	.36	.83	.52	.47	.57	.78	.53
H+W 3	M	3.62	2.58	1.60	1.18	2.42	2.16	1.72	1.38
	SD	.65	.83	.50	.16	.90	.28	.53	.51
H+W 4	M	3.73	2.76	2.00	1.60	2.68	2.42	2.06	1.53
	SD	.61	.95	.79	.49	.77	.29	.72	.58
H+W 5	M	3.96	3.24	1.47	1.29	2.70	2.49	2.06	1.47
	SD	.47	.52	.33	.25	.73	.62	.82	.46
H+W Mean	M	3.72	2.99	1.90	1.53	2.77	2.50	2.11	1.60
	SD	.42	.53	.45	.28	.58	.30	.63	.47
Liking	M	4.18	3.20	2.83	2.11	3.30	3.11	2.30	2.22
	SD	.60	1.13	.72	.78	.48	1.05	1.06	.83
Affect	M	3.53	3.32	2.81	3.11	3.38	3.30	3.37	2.46
	SD	.81	.59	.83	.92	.58	.56	.60	.48
Self - Esteem	M	3.81	3.29	3.78	3.44	3.66	3.58	3.21	3.46
	SD	.39	.44	.31	.64	.43	.26	.78	.73
		n = 11	n = 10	n = 12	n = 9	n = 9	n = 10	n = 9	n = 10

Table 2

Lecture 2 (L1) X Lecture 2 (L2) X Instructions (I)
Analyses of Variance for Seven Rating Measures,
Liking, Affect and Self-Esteem

		L1	L2	I	L1 X L2	L1 X I	L2 X I	L1 X L2 X I
Overall	F	n.s.	105.63	7.59	n.s.	16.74	12.40	n.s.
	p		< .0001	.007		.0001	.0007	
	ω^2		.47	.03		.07	.05	
H+W 1	F	n.s.	63.60	4.33	n.s.	4.97	5.92	n.s.
	p		< .0001	.0411		.0288	.0174	
	ω^2		.40	.02		.03	.03	
H+W 2	F	n.s.	83.38	n.s.	n.s.	19.72	5.61	n.s.
	p		< .0001			< .0001	.0206	
	ω^2		.46			.10	.02	
H+W 3	F	n.s.	82.47	5.83	n.s.	14.69	12.97	n.s.
	p		< .0001	.0183		.0003	.0006	
	ω^2		.42	.02		.07	.06	
H+W 4	F	n.s.	50.59	5.07	n.s.	12.12	4.97	n.s.
	p		< .0001	.0274		.0009	.0289	
	ω^2		.33	.03		.07	.03	
H+W 5	F	n.s.	150.50	6.23	n.s.	11.73	31.28	n.s.
	p		< .0001	.0148		.001	< .0001	
	ω^2		.55	.02		.04	.11	
H+W Mean	F	n.s.	128.64	7.33	n.s.	19.21	16.05	n.s.
	p		< .0001	.0084		< .0001	.0001	
	ω^2		.52	.03		.07	.06	
Liking	F	n.s.	31.82	n.s.	n.s.	6.60	n.s.	n.s.
	p		< .0001			.0122		
	ω^2		.26			.05		
Affect	F	n.s.	8.12	n.s.	4.57	n.s.	n.s.	n.s.
	p		.0057		.0359			
	ω^2		.08		.04			
Self-Esteem	F	4.90	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	p	.030						
	ω^2	.05						

Note: Degrees of freedom for each effect: 1,72.

n.s. = not significant

Table 3

Lecture 1 (L1) X Lecture 2 (L2)
Analyses of Variance at Each Level of Instructions
for Seven Rating Measures and Liking

		L1		L2		L1 X L2	
		Lecture 2 Instructions	Lecture 1 & 2 Instructions	Lecture 2 Instructions	Lecture 1 & 2 Instructions	Lecture 2 Instructions	Lecture 1 & 2 Instructions
Overall	F	12.55	4.97	88.48	25.45	n.s.	n.s.
	p	.0011	.0325	< .0001	< .0001		
	ω^2	.08	.06	.61	.37		
H+W 1	F	n.s.	n.s.	50.09	17.25	n.s.	n.s.
	p			< .0001	.0002		
	ω^2			.53	.29		
H+W 2	F	9.07	10.52	75.12	21.26	n.s.	n.s.
	p	.0046	.0026	< .0001	.0001		
	ω^2	.06	.14	.60	.30		
H+W 3	F	15.62	n.s.	85.66	14.10	n.s.	n.s.
	p	.0003		< .0001	.0006		
	ω^2	.10		.59	.25		
H+W 4	F	8.93	n.s.	39.82	13.67	n.s.	n.s.
	p	.0049		< .0001	.0008		
	ω^2	.09		.43	.24		
H+W 5	F	12.51	n.s.	304.61	14.16	4.59	n.s.
	p	.0011		< .0001	.0006	.0387	
	ω^2	.03		.84	.25		
H+W Mean	F	16.36	5.40	147.28	21.78	n.s.	n.s.
	p	.0002	.0263	< .0001	.0001		
	ω^2	.08	.07	.72	.33		
Liking	F	11.10	n.s.	22.71	10.78	n.s.	n.s.
	p	.0019		< .0001	.0024		
	ω^2	.14		.30	.21		

Note: Degrees of freedom for lecture 2 instructions: 1,38
Degrees of freedom for lecture 1 and 2 instructions: 1,34

n.s. = not significant

Table 4

Primacy and Recency Effects: Lecture 1 and Lecture 2
Means for Seven Rating Measures, Liking, Affect and Self-Esteem

	Primacy Effect				Recency Effect			
	Lecture 2 Instructions		Lecture 1 & 2 Instructions		Lecture 2 Instructions		Lecture 1 & 2 Instructions	
	Poor L1	Good L1	Poor L1	Good L1	Poor L2	Good L2	Poor L2	Good L2
ıll	2.74	2.01	1.77	2.20	1.40	3.35	1.51	2.46
	2.85	2.51	2.31	2.53	2.02	3.34	2.07	2.77
2	3.01	2.48	2.22	2.86	1.98	3.51	2.09	2.99
3	2.61	1.88	1.77	2.07	1.39	3.10	1.55	2.29
4	2.86	2.18	1.97	2.37	1.80	3.24	1.79	2.55
5	2.71	2.26	1.98	2.38	1.38	3.60	1.76	2.59
Mean	2.81	2.26	2.05	2.44	1.71	3.35	1.86	2.63
ng	3.50	2.66	2.66	2.80	2.47	3.69	2.26	3.20
ct	3.17	3.21	2.88	3.37	2.96	3.42	2.91	3.34
-Esteem	3.79	3.36	3.52	3.43	3.61	3.55	3.35	3.62
	n = 23	n = 19	n = 18	n = 20	n = 21	n = 21	n = 19	n = 19

Appendix - Relationships Among Measures

This study employed a total of 12 dependent measures: seven ratings, liking, affect, self-esteem and two achievement measures. The two achievement measures were principally manipulation checks on the effect of lecture 1 and lecture 2 quality. The relationships among these measures were examined in three ways: across experimental conditions, across subjects pooled from all conditions, and within experimental conditions.

1. Across Experimental Conditions

One purpose of this study was to determine whether affect and self-esteem mediate the relationship between lecture quality sequence and attraction. This purpose is addressed by examining the relationship between affect, self-esteem, liking, and ratings across sequence conditions. This is so because the relationship as proposed arises due to lecture quality sequence. Pooled correlations and within condition correlations may be due to other factors, as will be discussed. Thus, for each instructions condition, the order of lecture quality sequence for affect and self-esteem was compared to the order of sequences for ratings and liking. As discussed earlier, the order of sequences for affect partially corresponded to the order for ratings and liking, suggesting only weakly that affect may mediate between lecture quality sequence and ratings and liking. In neither instructions condition did the order of lecture quality sequence for self-esteem correspond to the order for ratings and liking. Thus, there was no evidence of a mediating role for self-esteem.

The ordering of sequences for ratings and liking was also exa-

mined. This assessment determined whether the variation in the experimental manipulations affected the seven ratings and liking similarly. As seen in Table 4, in each instruction condition, the order of lecture quality sequence for all ratings and for liking fully corresponded. This suggests that the seven ratings and liking were influenced in the same way by lecture quality sequence.

2. Pooled Correlations

Correlations for subjects pooled from all conditions were computed. This analysis assessed whether students responded similarly to the various dependent measures. For example, regardless of the source of variation in students' affect or ratings, did that affect correspond to their ratings? (All correlations were statistically tested at the .05 level of significance).

Overall high significant intercorrelations were found among the rating measures, $r(78) = .69$ to $.94$, $p < .05$). Correlations between ratings and liking were significant and fairly high ($r = .53$ to $.68$). Thus, students responded to all the ratings and liking similarly. Correlations betwen affect and ratings and between affect and liking were significant and moderate ($r = .28$ to $.40$). These correlations indicate that those students who gave the instructor low ratings had correspondingly low affect and vice versa, suggesting that affect may play a role in influencing ratings and liking. (Note that affect, ratings and liking vary in this case as a result of either experimental manipulations or student differences, thus the mediating role of affect is not addressed by this pooled correlation). Affect was also significantly correlated with quiz 1 scores ($r = .30$) and quiz 2 scores

(r = .36). Thus, academic achievement may influence student feelings. Quiz 2 scores also showed fairly high significant correlations with ratings and liking (r = .43 to .63). Recent student achievement thus may affect their ratings. Self-esteem was not significantly related to any other measure (r = .007 to .23).

3. Within-Condition Correlations

Within-condition correlations were computed for instructions, lecture 1 quality, lecture 2 quality, and lecture quality sequence. Within any experimental condition, the co-variance of two measures results from sources other than the experimental variation. For example a correlation between affect and liking within the good lecture 1 quality condition may be due to individual subject differences or to the impact of lecture 2 quality or the instruction manipulation, or any interaction of these factors, but not to lecture 1 quality. Differences in within-condition correlations between levels of the independent variables assessed whether the experimental manipulations affected the relationships among the dependent measures. For example, does the relationship between affect and liking differ when lecture 1 quality is good as compared to poor? The significance of the differences between correlations was tested using Fisher's Z$_r$ transformation at the .05 level of significance.

a. Correlations among dependent measures for lecture 2 and lecture 1 and 2 instructions.

Comparisons were made between all dependent measures for lecture 2 and lecture 1 and 2 instructions. Intercorrelations among the rating measures were significant and high for both instruction levels. How-

ever, the relationship between the overall item and the other rating items was consistently greater when the instructions were to consider lecture 2, r(40) = .75 to .91, than to consider lecture 1 and 2, r (36) = .63 to .76. Four of the six differences between correlations were significant. The overall rating item was thus not as representative of all rating measures when instructions were to consider both lectures, suggesting that students did not respond as similarly to the overall and specific rating items. Perhaps more of an effort was made to consider information from lecture 1 when asked for an overall rating than specific ratings. (Correlations of liking with overall and other ratings were generally similar for each level of instructions).

Correlations of quiz 2 with rating measures were consistently higher for lecture 2 instructions than for lecture 1 and 2 instructions. Four of the seven differences between correlations were significant. These findings are consistent with the proposition that academic performance plays a role in student ratings. When students were instructed to consider lecture 2, their ratings were highly related to their quiz 2 score (r = .55 to .75). However, when students considered both lectures their performance on quiz 1 was also an influence, which lessened the relationship between quiz 2 scores and ratings (r= .25 to .45). That quiz 1 score was a greater influence under lecture 1 and 2 instructions is also suggested by consistently higher correlations between quiz 1 and ratings for lecture 1 and 2 instructions (r= .16 to .48) than for lecture 2 instructions (r= -.03 to .16), although this difference was significant only for the correla-

tion of quiz 1 and Hildebrand & Wilson factor 2 (HW2, organization and clarity).

b. Correlations among dependent measures for good and poor lecture 1 quality.

Intercorrelations among the rating measures were significant and high for both good and poor lecture 1 quality (r = .65 to .95). Quiz 1 and ratings were significantly correlated when lecture 1 quality was poor, r (39) = .36 to .56; but not when lecture 1 quality was good, r (37) = -.01 to .16, suggesting that student performance on the lecture 1 quiz had more of an influence on ratings when it was poor than when it was good. This was found to occur, however, only when lecture 2 was good resulting in higher correlations for the poor-good lecture quality sequence as discussed in Section d. The differences between correlations for good and poor lecture quality reached statistical significance only for HW3 (instructor-group interaction).

Correlations between ratings and affect were moderate and significant when lecture 1 quality was poor (r = .37 to .58) whereas correlations were low when lecture 1 quality was good (r =.07 to .18). Five of the seven differences beween rating-affect correlations for poor and good lecture quality were significant. The higher correlations in the poor lecture 1 condition were primarily present when lecture 2 was good, as discussed in Sections c and d.

c. Correlations among dependent measures for good and poor lecture 2 quality.

Intercorrelations among the rating measures were moderate to high for both good and poor lecture 2 quality. Correlations between

overall rating and specific ratings were lower for poor lecture 2 quality, r (38) = .36 to .64, than for good lecture 2 quality, r (38 = .65 to .84, indicating that students gave overall ratings which were more similar to their specific ratings when they had just viewed a good lecture than a poor lecture. This suggests that overall and specific ratings provide different types of information when recent lecture quality is poor. Two of the six differences between correlations were significant.

Correlations between quiz 2 and ratings and liking were moderate when lecture 2 quality was good (r =.25 to .51) and low, negative and non-significant when lecture 2 quality was poor (r = -.06 to -.24). Five of the seven differences between correlations were statistically significant. This suggests that the previous finding that student academic performance plays a role in their ratings may be true only when recent lecture quality is good. This relationship is further qualified as correlations between quiz 2 scores and ratings and liking were higher when lecture 1 quality was poor, as discussed in Section d.

Quiz 2 scores also correlated significantly and moderately with affect when lecture 2 quality was good (r = .49), but the correlations was insignificant and low when lecture 2 quality was poor (r =-.049). This indicates that students' affect is influenced by academic performance only when recent lecture quality is good. Students who achieved high quiz scores when lecture 2 quality was good had high affect and those with lower scores had lower affect. When lecture 2 quality was poor, students' achievement was not related to their affect, perhaps

because they did not feel responsible for their achievement.

Affect was consistently and moderately correlated with ratings when lecture 2 was good (r = .29 to .49) but not correlated with ratings when lecture 2 was poor (r = .007 to .12). The differences between affect-rating correlations was significant only for HW3 (instructor-group interaction). Correlations were higher when lecture 2 quality was good primarily when lecture 1 quality was poor, resulting in higher correlations in the poor-good lecture quality sequence as discussed in Section d.

d. Correlations among dependent measures for poor-good, good--good, poor-poor, and good-poor lecture quality sequences.

A consistent trend was towards correlations among the rating measures and liking which were higher for the poor-good (PG) sequence (r (18) = .45 to .92) than the good-good (GG) sequence (r (18) = .40 to .92) and for the good-poor (GP) sequence (r (17) = .57 to .96) than the poor-poor (PP) sequence (r (19) = .13 to .87). This suggests a greater stability among ratings for the "gain" and "loss" conditions than for the invariant conditions. The change in lecture quality from either poor to good or good to poor may have forced a general evaluative decision about the instructor which guided students' responses to all rating items. Twenty-three of the 28 correlations were higher for the PG sequence than the GG sequence and 25 of the 28 correlations were higher for the GP sequence than the PP sequence, however, only seven of the differences reached statistical significance.

As noted in Section b and c, several correlations were found to differ depending on whether lecture 1 quality was good or poor

as well as whether lecture 2 quality was good or poor. These correlations may be viewed as differing as a function of lecture quality sequence.

Quiz 1 scores have been said to be more highly correlated with ratings when lecture 1 quality was poor. This occurred, however, only when lecture 2 quality was good. Quiz 2 scores have been said to be more highly correlated with ratings when lecture 2 quality was good. This relationship was greater when lecture 1 quality was poor. This combination results in quiz 1 and quiz 2 scores which were generally more highly correlated with ratings for the PG sequence (r = .26 to .69; .44 to .64) than the GG sequence (r = .04 to .41; -.05 to -.45). The relationship between student achievement and ratings is, hence, highest when previous lecture quality is poor and recent lecture quality is good. The difference between correlations of ratings and quiz 1 and quiz 2 did not reach statistical significance except for HW3. Students may view an instructor who improves in lecture quality as being particularly responsible for the quality of the lecture and thus base their ratings on how highly the instructor enabled them to achieve.

Affect was more highly related to ratings when lecture 1 quality was poor and lecture 2 quality was good. Four of the rating measures were significantly correlated with affect. Correlations between affect and ratings were consistently, although not statistically, higher for the PG sequence (r = .31 to .59) than the other sequences (r = .02 to .46). Correlations of affect and ratings were also generally higher for the GG sequence (r = .10 to .46) than the PP and GP se-

quences (r = .02 to .29) demonstrating more of a relationship when lecture 2 quality is good. It appears that when lecture 2 quality was good, there was a tendency for students' affect to be positively related to, and perhaps an influence on, ratings of the instructor. This was particularly so in the gain situation.

Liking and affect were significantly correlated only in the PP sequence. (r = .43). One explanation is that the previously discussed relationship between achievement and ratings is mediated through affect within the poor-good condition. Low correlations between affect and liking in the poor-good, good-good, and good-poor sequences suggest that affect is unrelated to liking of the instructor when at least one lecture is good. Since liking is the measure most often used in interpersonal attraction research, this finding directly contradicts the assertion that affect influences attraction except in the case of consistent punishment. (Note that this analysis does not relate to whether affect elicited by lecture quality sequence influences liking).