Journal of Personality and Social Psychology
1968, Vol. 10, No. 4, 317–340

PATTERN OF PERFORMANCE AND ABILITY ATTRIBUTION: AN UNEXPECTED PRIMACY EFFECT[1]

EDWARD E. JONES, LESLIE ROCK, KELLY G. SHAVER
GEORGE R. GOETHALS AND LAWRENCE M. WARD

Duke University

6 experiments were conducted to investigate the effects of varying distributions of success and failure on attributions of intellectual ability. In the 1st 4 experiments *Ss* confronted a stimulus person who solved 15 out of 30 problems in a random, descending, or ascending success pattern. In the 5th experiment only the descending and ascending patterns were compared. Contrary to prediction, the performer who showed improvement (ascending success) was not consistently judged to be more able than the performer with randomly spaced successes. The performer with a descending success rate, however, was consistently judged to be more intelligent and was expected to outperform those with either ascending or random patterns. Memory for past performance was uniformly distorted in favor of recalling more success for the descending performer and less success for the ascending and random performers. Neither this measure nor ratings of intelligence required, for their discriminating effects, that *S* himself solve the problems in parallel with the person being judged. In the final experiment *S* himself performed in an improving, deteriorating, or random but stable fashion, and was then asked to estimate his future performance. Under these circumstances, the ascending performer was more confident about his ability than the descending or random performer, reversing the picture of the 1st 5 experiments. The results were discussed in terms of the salience of early information in attributing ability and the role of social comparison processes.

Heider's (1958) analysis of person perception has generated considerable interest in the attribution process or in the conditions under which various personal dispositions are inferred from observed acts. Recent theoretical extensions of Heider's work by Jones and Davis (1965) and Kelley (1967) reflect this continuing interest. The number of empirical studies dealing with attribution is not large, and the bulk of these focus on such dispositions as motives, intentions, and attitudes. Ability dispositions have been almost entirely ignored in considering various attribution instances.

The present series of studies does focus on inferences about ability drawn from observed performance. There are undoubtedly many conditions that affect the extent to which ability can be clearly inferred from a given level of performance. The current studies concentrate especially on the role of various temporal patterns of success-failure over a series of intellectual problems; they ask whether more ability is attributed when a performer does increasingly well, increasingly poorly, or spreads his successes and failures evenly throughout the trials, assuming that the total number of correct solutions is held constant. The initial experiment was conducted to test the hypothesis that more ability will be attributed when the observed performance pattern shows a systematic trend (either up or down) than when the pattern is essentially random. The reasoning behind this hypothesis will shortly be presented. In addition, the authors were interested in the effects of trend direction itself, but here the theoretical guidelines were vague or conflicting. "Common sense," according to the clear consensus of those asked to predict the ability attribution outcome, favored the hunch that more problem-solving ability would be attributed to a performer showing

[1] This research was facilitated by National Science Foundation Grant G8857. The authors are indebted for comments and suggestions to Jack Brehm, Darwyn Linder, Roland Radloff, and John Thibaut. Catherine Bayes served as the experimenter in the first study and on film; she made a number of perceptive contributions to the planning of the first study and the authors are grateful for her help.

systematic improvement than to one whose pattern of successes showed a declining trend. The results were in the opposite direction. The remaining five experiments reported in this paper were attempts first to replicate and then to understand the reasons for this reversal of the common sense expectations.[2]

Dispositional Properties of Ability

The gist of Heider's reasoning about inferences concerning ability is that a given performance is a function of four factors: (*a*) environmental difficulty, (*b*) ability, (*c*) motivation, and (*d*) luck. The present experiments were conducted within a standard situation constructed so that environmental difficulty was high and constant and a relatively high level of performer motivation could be assumed. Thus ability and luck remained the two most prominent candidates for explaining a given level of performance, with ability being the more likely candidate to the extent that luck could be ruled out (and vice versa). Ability is considered here to include not only sheer intellectual prowess but also those stable temperamental factors that facilitate the fulfillment of intellectual potential—as distinct from more purely motivational variables such as "caring" or trying hard. Ability is thus used in a manner quite similar to Heider's (1958) concept of "power."

Under certain conditions of unrestricted casual observation, it is not difficult to distinguish ability from luck. If one observes consistent success on tasks below a certain known level of difficulty and consistent failure on tasks above that level, the question of luck need not even arise in attributing ability. There are, in addition, certain kinds of tests against nature that provide highly salient ability cues: lifting a 1,300-pound weight, running a mile in 3:51, swimming the English Channel. Finally, there are complex intellectual performances where the solution options are open-ended or multiple; whenever the chances of accidentally hitting upon the correct answer approach zero, the role of ability is correspondingly prominent.

The attributional problem of sorting out the relative contribution of luck and ability arises when there are complex interactions between personal and environmental factors (as in pinball) and only limited observation is possible, or when there are a small number of response options (as on a true-false or multiple-choice test). The present experiments deal with the solution of difficult progressions and analogies, each ending in four multiple-choice response options. A given correct answer can obviously, then, be either a true reflection of ability or no more than a lucky guess.

A further characteristic of ability—at least in the conceptual framework of the typical naive observer—is its relative permanence as a dispositional attribute. Ability does not come and go; only the conditions favorable to its manifestations do. Heider (1958) gives the example of an alcoholic or psychotic "who is able to hammer a nail, read a book, and so on only on a rare day of lucidity. We are led to feel that he can do these things even though there are few occasions when he can actually do them [p. 95]." An intriguing implication of this conceived durability of ability is that once one decides a person is or is not able, it may be customary to attribute variations in performance to nonability factors such as luck, motivation, neurosis, the envirenment, etc. This implication seems borne out by the results to be reported, and shall be dealt with later.

Given a process of attribution where the inferential route from performance to disposition is overlayed with amplifying and concealing "extraneous" factors, there is ample room for the perceiver's motives to affect the attribution outcome. A perceiver who "needs" to see a performer as lucky rather than able, or vice versa, can find ample supporting evidence in ambiguous settings similar to the present experimental paradigm. The nature of the perceiver's relationship to the performing target person is of particular importance as a motivational determinant because the attribution of ability of others is closely linked to the attribution of ability to oneself.

[2] The experiments were run roughly in the order that they are presented, but there was often considerable overlap in time so that the design of one study did not always have the benefit of results from the one that preceded it.

This conclusion is not only intuitively apparent, it is a cornerstone of Festinger's (1954) theory of social comparison. The most relevant propositions from Festinger's theoretical statement are those concerning what might be called the autistic control of ability attribution. Two motives are involved in the ability domain. We want to see ourselves as more capable than others and yet know *how* capable we really are with some precision. Biased attribution can occur in response to either or both motives. We can attribute to others less ability than they in fact possess, thus enhancing our own relative standing, or we can attribute to others an ability level more similar to our own than the facts justify, thus enabling more precise comparison.

When someone does better than we do, both of these tendencies would seem to operate in the same direction, to draw the person closer to us. The tendencies would seem to be in conflict when someone's performance is lower than our own. However, there are subtleties that complicate these propositions. Certainly people are aware that in most ability realms there are some who are truly more talented than they and others who are truly less so. Given sufficient evidence of noncomparability of performance, then, there might be clear restraints on any tendency to perceive another person as less able, more comparable, or both. In fact, there may arise motivational pressures to *reduce* the comparability between oneself and another person. When a performer clearly does better than we have done, we may actually protect our self-esteem by attributing unusual ability to such a person and thus rendering him noncomparable. If we can see the superior performer as a genius, a "pro," an unusual natural talent, this enables us to maintain our own self-esteem at or near its former level. This is the obverse of Hakmiller's (1966) proposition that threatened persons will seek out dissimilar, inferior persons with whom to compare, as a salve for their own self-esteem. The implications of this reasoning for a primacy effect in ability attribution will be further developed in discussing the expermental results.

Basic Experimental Setting

In each of the following experiments the same sets of analogies and progressions were used as performance-eliciting test items. The experiments differed, as shall be seen, in the form and context of item presentation. In the first five experiments the target or stimulus person (SP) was the same female performer attempting to solve two 30-item series of multiple-choice problems. In the sixth experiment male subjects and a male SP were involved. Two-thirds of the items in each series were insoluble; the remainder were sufficiently difficult so that the entire series of items was designed to appear demanding in the extreme. The SP attempted to solve the first 30 problems, and after each solution attempt the experimenter provided predesigned feedback on the correctness of the attempt. In the initial three experiments the subject herself attempted to solve the same 30 problems along with the SP. She also received predesigned feedback. In the major experimental conditions the SP always "solved" 15 out of 30 problems and the subject always "solved" 10—according to the feedback delivered.

On the second series of 30 problems the roles of the subject and the SP diverged. Presumably by chance, the SP was instructed to continue attempting to solve each successive problem in the time allotted. The subject was instructed to try to predict after each solution attempt whether the SP's answer was correct or incorrect. Since the subject was given no information about which answer the SP chose and there was no feedback during the second series of 30 items, the resulting prediction should have provided a relatively direct measure of the SP's perceived power over the task and at least an indirect measure of ability attribution. Additional measures of performance recall and rated intelligence were also obtained.

In all but one experiment the basic comparison was between three patterns of SP performance: descending success, ascending success, and random success. (The last pattern was omitted in Experiment V.) As indicated above, the prediction leading to the first experiment was that the two systematic trend conditions would lead to greater perceived power over the task than the unsystematic condition. This prediction assumes a general belief that chance or luck—as reflected in "guessing"—would produce sporadic successes in an unsystematic pattern. Such a belief would in fact be in tune with the true average probabilities. A more psychologically interesting way to put it might be to say that evidence of control over the task at any point implies the ability, in general, to control such tasks. During those phases of the task where performance is low, this ability is being obscured by motivational and other factors.

The difference between ascending and descending patterns in ability attribution was not readily predictable for existing theory and was approached at the outset as an empirical question. It is rather easy to list comparable reasons favoring the attribution of superior ability either to the ascending SP or to the descending SP. In the ascending case perhaps (*a*) the SP took a while to catch on, but having done so he would continue to perform well; (*b*) the SP realized he was doing poorly and began to try harder, pay closer attention, etc.; (*c*) recent events (in this case predominant success) were better remembered by the perceiver than more remote ones; (*d*) having initially underestimated the SP's ability the perceiver "overcompensated" when he learned that his initial judgment was too severe (cf. Walster, Walster, Abrahams, & Brown, 1966). In the descending case perhaps (*a*) as often is the case, the earlier problems were easier; (*b*) the SP realized that he could solve the problems and lost interest; (*c*) a primacy or "prior entry" effect, observed in other first-impression contexts, also operates in the ability realm; (*d*) having made an initial judgment to which he feels committed, the perceiver reduces the dissonance created by disconfirming subsequent information by ignoring or devaluing it (Walster & Presthold, 1966).

Experimental procedures were designed to minimize the role of the first factor in both cases. "Catching on" was not a salient possibility in view of the discrete, uncorrelated nature of the problems plus the fact that similar problems are common to many intelligence and aptitude tests and familiar to most college students. It was difficult to conclude that the early problems were easier because all subjects were explicitly told that the problems were equal to each other in normative difficulty; in addition, each subject received bogus feedback indicating that she did no better on early than on late problems. The first experiment would supposedly focus, then, on the relative importance of primacy versus recency of information about successful performance, on vagaries in the attribution of motivation for declining versus improving performance, and on subtle factors in the realm of cognitive overcompensation or dissonance reduction. The experiment was directed more toward finding out whether anything needed to be explained than toward providing a clear explanation of observed differences.

EXPERIMENT I

Method

Subjects and basic design. A total of 43 female subjects, recruited from an introductory psychology course, served as perceivers (predictors) in the experiment. One of these subjects was used in a practice session, one was eliminated because of experimenter error, and one was convinced that the instructions were deceptive. Each of the remaining 40 subjects was randomly assigned to one of five conditions. Ten were assigned to each of the three major conditions: ascending feedback pattern, descending feedback pattern, and random feedback pattern. In each of these conditions the SP (an accomplice of the experimenter) reportedly solved 15 of the first 30 problems correctly. The two remaining conditions, random high and random low, were included to examine the attributional effects of clear differences in overall performance. In the random high condition the SP allegedly solved 19 problems; in the random low condition she solved only 5. In each case there was no particular trend to the pattern of correct answers. There were only five subjects in each of these latter conditions.

Procedure. Each subject was introduced to the SP, who appeared to be another subject from a different introductory psychology section. The experiment was vaguely described as a study of "performance under social conditions," and the subject and the accomplice were instructed to work independently, but simultaneously, on the same problems. They were forewarned that all problems were taken from a larger battery of problems designed to discriminate at the very highest levels of intelligence and, therefore, they should not be surprised at their failure to obtain a large number of correct answers,

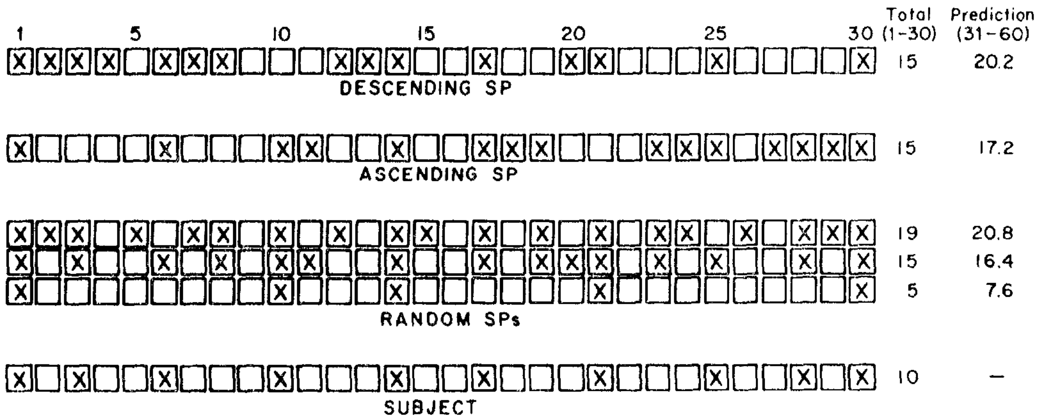

FIG. 1. Feedback patterns employed in first and, selectively, in other four experiments (Each box indicates a trial and each check indicates a success. Mean prediction scores for second series are listed in final column.)

At the beginning of a trial each subject was given a card displaying the same problem and was instructed to record her answer when signaled to do so (after 20 seconds). The experimenter (a female senior) tried to assure the subjects that she did not wish to put them under severe time pressure, but the time limit was devised so that the experiment would be completed under standard conditions and within the allotted time.

As the subject and the accomplice proceeded to work on the initial series of items, the experimenter examined their answers after each item and publicly informed each in turn that she had obtained the correct answer or that she was wrong. In this manner the predesigned variations in performance pattern were introduced. While the subject saw that the accomplice was increasingly, decreasingly, or sporadically correct, she herself received feedback indicating 10 randomly dispersed successes over the 30 items. In order to make the developing patterns salient, the subjects recorded successes in vivid "Magic Marker" stripes on a horizontal record sheet. The patterns employed for the subject and the SP are illustrated in Figure 1.

After the first 30 items were completed, the experimenter explained that she was really interested in studying one person's ability to predict the performance of another, after both had comparable experience with the performance task. The practical importance of predicting success on the basis of limited information was stressed. The experimenter then proceeded to assign two roles for the second part of the experiment. Allegedly on alphabetical grounds, the accomplice was assigned to be the performer and the naive subject was assigned to be the predictor. The experimenter announced that the subject whose name appeared earlier in the alphabet would be the predictor, or that she would be the performer, depending on the actual alphabetic priority of the subject's last name. Having assigned the two roles, the experimenter announced that the performer would receive $.25 for each correct answer and would lose $.10 for each incorrect attempt. The predictor, similarly, would receive $.25 for each accurate prediction and would be docked $.10 for each mistake. It was explained that the second series would consist of items of equal difficulty taken from the same general pool as items in the first series. The predictor would see each item in turn, be able to observe the performer trying to solve it, and at a stop signal each would record the appropriate answer or prediction. Feedback was not provided at any time during the second series. The accomplice made a strenuous effort to maintain inscrutability, and never recorded her answer to any item on either series until given the signal to do so. In this way possible cues from response latency were eliminated or minimized.

At the conclusion of the second, prediction series, the accomplice was taken to a separate room, and the subject filled out a questionnaire containing measures of recalled performance and ratings of intelligence. After the customary probes for suspicion, the various deceptions were revealed to the subject, she was reintroduced to the accomplice, and the true purpose of the experiment was discussed in detail.

Performance items. The progressions and analogies were largely invented for the occasion, though several were adapted from existing items. Approximately one-third of the items on each series did have a reasonable solution (e.g., Tam, tan; rib, rid; rat, raw; hip, ___ 1. hid, 2. hit, 3. his, 4. him), whereas the remainder did not (e.g., quality is to benevolence as judgment is to ___ 1. praise, 2. endearment, 3. criticism, 4. malevolence). The items were so ordered that the soluble ones appeared in those positions where the subject received "correct" as feedback. This was done to increase the plausibility of the feedback.

In every condition the items appeared in the same order within a series. An attempt was made to counterbalance the series order by randomly assigning some subjects to one order (Series 1 followed by Series 2) and the remaining subjects to the other order (2 followed by 1). Unfortunately, 8 out of 10 subjects in the random-pattern condition were thereby exposed to the 1-2 order, though an approximate 50-50 split was maintained in the other experimental groups. On the basis of circumstantial evidence from the first experiment and more direct evidence from the remaining experiments, this confounding of order with performance pattern is not considered to represent a serious problem. The fact that subjects were assigned an order on a random rather than a systematic or alternating basis was nevertheless an error in procedure.

Results

Validation of procedural variations and controls. The subjects apparently accepted the experimenter's assurances that the items were of equal difficulty. Thirty-one of the 40 subjects perceived the items in Series 1 as equally difficult; 6 (scattered across the different conditions) thought the early items might have been somewhat easier, while 3 thought the later items were. There was a tendency to see the second series as somewhat easier than the first, regardless of series order. Of those 21 subjects who noted a difference, 16 rated the second-appearing series as easier than that which appeared first. The meaning of this trend is not clear, but it probably has something to do with the fact that the subjects themselves attempted to solve the first series (with limited success) but not the second.

How clearly did the subjects perceive the pattern of performance to which they were exposed? All but five of the subjects checked the appropriate response category for their condition, indicating that the SP did better at first, better at the end, or about the same throughout. All but five of the subjects also indicated that there was no discernible pattern to their own successes.[3]

Measures of ability attribution. In predicting the SP's performance on the second series of 30 trials, the subject could make use of several kinds of information. She could weigh the difficulty of the item, read any signs of puzzlement or enlightenment appearing on the SP's face, note other signs of effort or ease of solution reflected in the SP's behavior, and finally the subject could shade her predictive decisions toward success or failure in line with her general impression of the SP's ability. If the SP fulfilled her assignment well, the summary prediction score should have reflected primarily the latter ability judgment. At least there should have been no differential emphasis on the other factors from condition to condition, whereas it was hoped that different patterns of performance would create different impressions of ability. Since there were no discernible patterns or gradients in the predictions themselves—subjects in the ascending condition, say, did not predict that the SP would be better at the end than at the beginning of the second series—the sum of predicted successes could serve as an adequate general measure of attributed ability.

The resulting means for each condition appear in the far right column of Figure 1. Not surprisingly, subjects in the high random condition predicted the greatest number of successes (20.8), and those in the low random condition predicted the smallest number (7.6). This makes it quite clear that subjects will base predictions of ongoing performance on the level of past performance when there are no other salient cues about the determinants of that performance—when successes are randomly scattered. The major interest, of course, is in the effects of the distribution of successes given in a standard level of overall performance. When a comparison is made among the three conditions in which the SP solved 15 problems, but in different patterns, it is evident that the subjects in the descending condition predict a higher level of success on the second series than subjects in either the ascending or the random-15 condition. The overall F ratio ($df = 2/27$) is 7.06, $p < .01$; orthogonal comparisons tests show that the descending condition is significantly higher than either the ascending or the random-15 condition, whereas the latter two do not significantly differ.

The most confident experimental hypothesis was that more ability would be attributed to both the ascending and descending SPs than to the random SP. This hypothesis is con-

[3] Unless otherwise noted, these results on "validation of manipulations" held for the subsequent experiments as well and will not receive further attention.

firmed when the two "systematic trend" conditions are pooled and compared with the random condition ($F = 6.20$, $df = 1/27$, $p < .05$). However, the unexpectedly high value of the descending condition is the main contributor to this confirmation, and the perceived superiority of the descending over the ascending SP provides the most obvious challenge to explanation. The average subject in the descending condition predicted that the SP would solve 20 of the 30 problems in the second series, even though the same person only solved 15 in the first series and ended up doing very poorly. This is about the same prediction level as that advanced for the random SP who solved 19 out of the first 30.

The prediction of performance on the second series is, of course, not a pure measure of perceived intelligence. It is presumably a measure of intelligence plus motivation plus various temperamental factors enabling the person to function well under moderate time pressure. Could the differences in prediction reflect differences in attribution motivation rather than ability? There is no support for this conjecture in the rating data. Subjects were asked to evaluate the level of concentration of the SP and her concern with doing well. There were no differences between conditions on these measures that even approached significance. Subjects were also asked to evaluate the SP's intelligence on a simple scale ranging from "well below average Duke student" to "well above average Duke student." They were also asked to evaluate their own intelligence on the same scale. There were no differences in intelligence attributed to the SP, but differences did emerge when the discrepancy between own and SP intelligence was recorded. All but 1 of the 30 subjects in the three major conditions rated the SP as more intelligent than herself, but for subjects in the descending condition the discrepancy was significantly greater than for subjects in either the ascending ($p < .01$) or random ($p < .10$) condition. This provides at least some evidence for construing the prediction of performance as a measure of ability attribution.

Recall of performance level. In the attempt to understand the high level of ability attributed to the descending SP, it obviously is relevant to note whether there is a corresponding distortion of recalled success. Each subject was asked, on the final questionnaire, "Approximately how many questions (out of 30) did your partner answer correctly in the first game session?" The average recall score of subjects in the descending condition was 17.8, in the ascending condition, 15.2, and in the random condition, 16.9. Scores of both the descending and random SPs were distorted upward, therefore, and the recalled success for the descending SP was significantly greater than for the ascending SP ($t = 2.11$, $df = 18$, $p < .05$).

Discussion

The results of the first experiment may be summarized as follows: Considering only those conditions in which the SP ended up with 15 out of 30 items correct, the descending SP was perceived to do better on the second series than either the ascending SP or the random SP. She was also judged to be more intelligent than either the ascending or random SP (relative to the subjects' self-ratings), and the subjects recalled the descending SP as actually getting a higher number correct on the initial, feedback series than they recalled for the ascending SP or the random SP, though the latter comparison was not significant. On each of these measures, then, the descending SP is seen to stand out from either the ascending SP, the random SP, or both.

How might the unexpected superiority of the descending SP be explained? One clue may lie in the abrupt personal significance of the subject's relative failure on the early trials. Although competition was not stressed, it was obviously inherent in a situation in which two "subjects" were getting trial-by-trial feedback about their performance. It is hardly far-fetched to imagine the naive subject in a state of initial apprehension concerning her relative performance on a "high level intelligence test." Since she was told that the items were very difficult, she would be less concerned with her absolute level of success than with her success in comparison to a fellow student. In the descending condition, therefore, there is an apprehensive subject whose fears are rapidly confirmed: her partner solves seven of the first eight problems correctly, while she can manage to solve only three herself. Perhaps the most attractive way for the subject to handle such a development is that which was mentioned in the introduction. The subject can best maintain her own self-esteem by rendering the partner noncomparable and seeing her as unusually talented. Because she has ceased comparison it is not difficult for the subject in the descending condition to distort her recall in the direction of her initial impression. In the contrasting ascending condition, however, the subject is enticed into an initially comfortable, competitive relationship with the partner, and by the time the partner begins to pull ahead in performance the subject is already committed to a continuing ability comparison. The result: accurate recall of the general level of success in the ascending condition. Subjects in the random condition recall a somewhat higher performance on the first series than subjects in the ascending condition, but predict a lower performance on the second series. This is fairly understandable if the SP's initial performance level is considered to reflect a fair number of lucky guesses.

Since the superiority of the descending SP was an unexpected finding, such speculations are really premature unless the differences observed in the first experiment can be replicated under more rigorous conditions. Not only might the findings be a quirk of chance; the accomplice could have unwittingly influenced the subject's judgments of ability. It was impossible to conceal from her the condition to which the subject was assigned, and she might have emitted systematically different cues in the different conditions. In addition, the partial confounding of condition with series sequence has already been noted, raising the possibility that the descending SP might have been judged more able because the particular items she solved were seen as more difficult. Experiment II was designed in an effort to remove these sources of artifact and to see if the unexpected superiority of the descending SP would again obtain. The second experiment was also designed to examine the effects of the subject's own performance on her judgments of SP's ability.

EXPERIMENTS II AND III

Method: Experiment II

In preparation for the second, third, and subsequent experiments a sound film was made that featured the the same accomplice attempting to solve the same problems used in Experiment I. In Experiment II, groups of subjects viewed the film with an appropriate introduction, attempted to solve the problems along with the filmed SP, made predictions of the SP's performance on a second series of problems, and answered a final questionnaire. Feedback concerning the correct alternative and the performance of the SP was given on the first problem series.

Stage-setting instructions. Subjects were immediately told that the experiment was concerned with the prediction of intellectual performance and was an extension of a study done the previous semester. The gist of the previous experiment was explained to them along with the desire to streamline and simplify the procedure for use with larger groups of subjects. Each group of subjects, they were told, would be shown a film of a different experimental session. This comment was designed to reassure them that there was nothing necessarily special or atypical about the subject they were to observe. It was added that the session was filmed with a hidden camera but that the filmed subjects had later given their permission to use the film in a follow-up study. The actual subjects were then told that they would be asked to solve problems along with the subjects on film merely to familiarize themselves with the problems so that information about the successes and failures of the filmed subject would have more meaning. They were specifically told not to put their names on their answer sheets, but were forewarned that they would be asked to identify their experimental materials by number before handing them in. The first part of film was then shown.

Presentation of the film. The film involved three players: the original female undergraduate experimenter, the undergraduate accomplice (seated facing the camera), and a third female whose back was always to the camera. Part 1 of the film began with instructions to the filmed subjects identical to those used in the original study. The film proceeded through the point at which the two "subjects" handed the experimenter their solutions to the first problem and were told they both were correct (see Figure 1). At this point the live experimenter said,

> We found that last year's subjects had a great deal of trouble trying to do much of anything besides simply solve the problems. Consequently, we are not going to ask you to watch the film at the same time you are working on the problems.

Each of the remaining 29 problems of the first series was then projected on the screen for 20 seconds. As each problem faded from view, the subjects were to record their answers. The correct alternative for that problem was then announced, and the experimenter went on to report whether or not the filmed subject had obtained the correct answer. This was further indicated by a checkmark in an appropriately numbered box on the blackboard. In this way the descending, ascending, and random conditions were created, and the sequence of checked and unchecked boxes remained on the board until just before the final questionnaire. Thus the pattern of performance was at least in the visual background during the prediction task.

It was obviously not possible to cite a particular alternative as correct and also control precisely the scores obtained by the naive subjects. To further the objective of keeping the subjects' totals as close as possible to the score of 10 that each was assigned in the first experiment, correct answers were given for the 10 soluble problems (which fell, it will be recalled, in a random pattern), and the answer judged to be *least* likely was chosen as correct for the remaining 20 problems.

After the various conditions had been established by predetermined feedback, the subjects were instructed to observe the second portion of the film and to attempt to predict whether the filmed subject was correct or incorrect for each problem in a second series. The fact that the predicting subject in the original experiment was paid for her correct predictions was mentioned, and the experimenter apologized that he could not pay subjects in the present experiment. The second portion of the film then began. In it the SP was shown attempting to solve each of 30 problems, the experimenter (on film) allowing her 20 seconds for each. The SP, as per instructions, always waited until the 20 seconds were up before recording her answer. The subjects watching the film were not only exposed to the SP's (hopefully, unrevealing) face, but also the particular problem was itself overdubbed toward the bottom of the film so that the SP and the problem could be simultaneously viewed. Each subject also had a packet of problems in front of her in case she wanted to study them more intensively. After the film was over, the subjects filled out a final questionnaire and the deceptions and purposes of the experiment were fully disclosed to them.

Design and subjects. Ninety-five female subjects were again drawn from the introductory psychology course in the semester following the one in which the first experiment was run. These subjects were run in six groups of 19, 19, 19, 17, 12, and 9. As a partial corrective for the confounding of sessions and conditions, two sessions each were run to compose the ascending and descending groups. The means for all major dependent variables were also identical across sessions within a condition. Also, the experimenter was unaware of the condition he was running until the point at which feedback was delivered, except at the last session. Two random groups were included in the design. One of these was included to control for the content of those items designated correct in the various conditions. While the first random group was comparable to the random condition of the first experiment, the second random group received the items of the first series in such an order that the SP was correct on exactly the same items (but still in the random pattern portrayed in Figure 1) as the descending SP. If the results of the first experiment were in some way contingent on the particular items being solved, then the superiority of the descending SP should not be maintained over the random SP who solves the same items but does so in a different order.

Results: Experiment II

Subjects' own scores: Reconstituting the sample. The authors' attempt to control roughly the subjects' own scores was less successful than had been hoped. Perhaps by combining their ability to solve most of the soluble problems with good luck on the remaining problems, subjects on the average were correct more than a third of the time (11.5). Number of items correct ranged from 5 to 19, raising obvious problems of comparison with the initial experiment. These problems would be acute if there were a substantial correlation between one's own score and one's predicition of the SP's score. This correlation was negligible with the total sample ($r = -.12$), and within cell correlations did not differ significantly from each other.

Unexpectedly, there were significant variations in average "own score" across experimental groups ($F = 6.12$, $df = 3/85$, $p < .01$). Differences between the two random conditions were largely responsible for this overall effect. Subjects in the first random condition (run as the first experimental group) averaged 12.89 items correct, whereas those in the second random condition (run last) averaged only 9.33 correct. It is very unlikely ($p < .001$) that this is merely a chance difference and there are a number of factors that could be responsible: any factor associated with time during the semester or the position among groups run, any factor associated with unwitting differences in the experimenter's behavior, or any factor associated with item order. Apropos of the last factor, it will be recalled that the two randdom conditions presented the items in different orders such that the SP in the second random condition was given "correct" as the feedback for the same items as the descending SP. However, the subjects solved the same 30 items in all conditions and received the same feedback. It is difficult to see, therefore, why the order of items should be important.

Although items correct for the subject did not correlate systematically with predictions for the SP on the second series, it was decided to equalize the conditions with respect to "own scores" by eliminating all subjects who scored more than 13 or less than 7 correct. This reduced the total sample size to 64 and the resulting "own score" variance attributable to conditions became negligible ($F = 1.11$, $df = 3/60$).

Ability attribution. Once again subjects in the descending condition predicted a high level of SP performance on the second series of 30 items. As Table 1 shows, the mean prediction score for the descending condition was significantly higher than scores for the ascending and random conditions, the latter being quite similar to each other. The same pattern of differences obtained on the questionnaire ratings of intelligence. There are slight differences between the two random conditions on these two measures, but their direction is not consistent. The descending condition mean is closer to random-1 on the rated intelligence score and closer to random-2 on the prediction summary score. With some slight residual trepidation, then, it was concluded that the differences in attributed ability observed in this and the previous experiment were not a function of the particular items allegedly solved by the SP. The pattern of successes itself seems to be the critical thing. It might be noted, finally, that the results remain essentially the same when the total rather than the reduced sample is used.

TABLE 1

VARIATIONS IN ABILITY ATTRIBUTION: EXPERIMENT II

Condition	n	Prediction summary[a]	Rated intelligence[a]	Recall SP performance[a]
Descending	31	18.65	7.50	16.79
Ascending	19	15.79**	6.35**	14.65**
Random 1	9	15.56*	7.00	14.89*
Random 2	8	16.50[b]	6.63*	14.63[b]
Random 1 + 2	17	16.00*	6.82*	14.76*

[a] All comparisons indicated by asterisks are between the descending condition and the one indicated. Comparisons are based on t tests.

[b] These means differed from respective descending means at the .10 level.

* $p < .05$.

** $p < .01$.

Recall of performance level. On the final questionnaire subjects were asked, as in Experiment I, to recall the number of questions answered correctly by the SP on the first series. The final column in Table 1 shows that the descending SP again stands out. Subjects in this condition recalled that the SP was more often correct than she in fact was and more than the comparable recall scores for the other conditions. The ascending and random conditions obviously did not differ among themselves.

Method: Experiment III

In an additional experiment, not to be fully reported here, 67 female subjects were exposed to the same feedback sequences of ascending, descending, or random success that were charted on the blackboard in Experiment II. Subjects were run four at a time. Each subject sat in a semi-enclosed booth containing electronic signaling devices that gave her standard feedback concerning her own performance from problem to problem. Although subjects were led to believe the feedback was individualized, they each received the same pattern of 10 rights and wrongs that was devised for Experiment I. By instructional variations, some subjects were led to believe their performances would be publicized, whereas others were assured that only they would know the level of their own success. There was thus a total of six experimental conditions in a 2 (public versus private) $\times$ 3 (ascending, descending, or random) factorial design.

Results: Experiment III

There were no discernible effects of the private-public variation. In both private and public conditions, the descending SP was again judged to be more intelligent than the ascending SP ($t = 2.92$, $p < .01$). She was also recalled as having solved more of the first series of 30 problems than either the random SP ($t = 4.59$, $p < .001$) or the ascending SP ($t = 4.21$, $p < .001$). Trial-by-trial predictions of SP success on the second series of problems were not significantly affected by the pattern of success on the first 30 trials. However, when asked to return to the second 30 problems and to estimate whether they would have solved each one in turn, subjects in the ascending condition thought that they would have done better than subjects in the descending and random conditions. The discrepancy between the level of success the subjects attributed to the SP and to themselves was therefore smaller in the ascending than in the descending condition ($t = 2.47$, $p < .05$). The means for the crucial comparisons of Experiment III are presented in Table 9, which summarizes the results of the first five experiments.

Correlations between self-estimates and predictions for the SP. The fact that subjects in the ascending condition estimated their hypothetical performance as more similar to their prediction for the SP than subjects in the remaining conditions suggests that subjects confronting the ascending SP remain more competitive throughout the experiment and find it more difficult than subjects in the other conditions to concede superiority to her. To delve further into this problem, the subjects' self-estimates were correlated with their predictions for SP in each of the six conditions. The results are presented in Table 2. They show that subjects in the ascending and random conditions seem most intent on maintaining comparability, whereas descending-condition subjects judge the SP independently of their self-estimated capability. This pattern of correlations is consistent with the hypothesis of comparison cessation in the descending condition.

TABLE 2

CORRELATION BETWEEN ESTIMATES OF HYPOTHETICAL OWN SCORE AND PREDICTION OF SP SCORE ON SECOND SERIES: EXPERIMENT III

Condition	n	Product-moment correlation
Descending		
Private	12	.02
Public	10	.43
Ascending		
Private	12	.81**
Public	10	.89**
Random		
Private	11	.64*
Public	12	.73*

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

Discussion: Experiments II and III

The major contribution of the second experiment was to replicate the distinctiveness of the descending condition on the three measures—predicted performance, rated intelligence, and recalled performance. Since the SP was on film and her expressions and behavior were identical in all conditions, it would be difficult to attribute the results of Experiment I to unwitting influence by the accomplice. The fact that the two random groups made very similar predictions and ratings also tends to rule out item content as a factor responsible for the differences in attributed ability.

Experiments II and III both state that a face-to-face confrontation is not essential to the results of attributing superiority to the descending SP. Since the subjects' own scores were variable and anonymous in Experiment II, the result does not depend on public competition and correlated face-saving maneuvers. Consistent with this, the absence of any discernible effects of the public-private variation in Experiment III raises doubts about the importance of face-saving and defensively motivated cessation of comparison in the descending condition. If the maintenance of self-esteem were an important contributor to the observed primacy effect, one would surely expect the effect to be amplified by the public and muted by the private condition.

We still do not know, however, whether the primacy effect holds when the subjects make no attempt themselves to solve the problems. Even in the private condition the subjects can, of course, compete with the filmed SP, and discrepancies in subject-SP performance can have meaningful implications for self-esteem. One bit of correlational evidence has been unveiled suggesting that subjects in the ascending condition are more highly motivated to maintain comparison with the SP than subjects in the descending condition. It is difficult, therefore, to relinquish completely the notion that some of the observed differences reflect esteem-maintaining cognitive maneuvers.

But, on the other hand, suppose that the results reported thus far can be parsimoniously explained as a durable perceptual or cognitive illusion, that people do not necessarily *want* to see the descending performer as smarter than the ascending performer as a cushion to their self-esteem but are drawn into errors of performance recall as fatally as they are drawn to error in the Müller-Lyer illusion. What might be the basis for such "unmotivated" cognitive distortion?

One interesting possibility is that subjects make continuously shifting probability-of-success estimates in observing the SP's performance, and are heavily influenced by these estimates in their subsequent recall attempts. Following Atkinson's (1964, p. 258) suggestion that "the strength of expectancy of success . . . is a function of the number of times the act has been followed by success divided by the number of times the act has been performed," it is possible to assign a subjective probability value for each trial. That is, the attending subject may calculate the subjective probability of success on the next trial from his knowledge of success on preceding trials, both for himself and for the SP. If he were to do this, even in some quite implicit or very crude way, there would be dramatic differences between the various patterns in the unfolding of these probabilities. Obviously, since the descending and ascending SP both end up solving 15 out of 30 problems correctly, the final subjective probability value should be .50 in each case. For all the trials leading up to this final point, however, the probability of success on the next trial is always higher for the descending than the ascending SP. This means that the subject must confront and adjust to a prolonged condition of SP superiority in the descending condition.

These claims are reflected in Table 3, which presents the theoretical values for SP and self, summarized by blocks of 10 trials. Here it is clear that the subjective probability scores are not only higher at all points in the descending SP conditions, but the discrepancy between the subject's own success expectations and his predictions of SP success follow the same trend. By comparison, there is a negligible discrepancy in the ascending condition throughout the series, in spite of the fact that the subject eventually solves five less problems than the SP. The high level of success on the early trials generates very high probability values. In the ascending pattern the identically high level of success on later trials does not affect the probability judgments so dramatically because the later estimates are heavily weighted by what has gone before.

TABLE 3

MEAN (THEORETICAL) SUBJECTIVE PROBABILITY ESTIMATES FOR SELF AND SP ON TRIAL ($n + 1$) BASED ON NUMBER OF SUCCESSES DIVIDED BY n

Trial block	Descending			Ascending		Random	
	Self	SP	SP-Self	SP	SP-Self	SP	SP-Self
2–10	.46	.87	.39	.30	−.16	.49	.03
11–20	.33	.65	.32	.36	.03	.48	.15
21–30	.31	.54	.23	.43	.12	.50	.19
Total	.37	.69	.32	.36	−.01	.49	.12

The hypothetical numbers in Table 3 do not, of course, prove anything, but they do show how a concern with subjective probability might set the stage for distortion in recalling the level of performance. In order *not* to distort, the subject must restrict his attention to the very last probability estimate in the entire cumulative series. If one realizes the rather special nature of ability as a dispositional attribute, one can see why such a rational restriction of attention might be thwarted. First of all, for the subject who is concerned with social comparison, the history of an ability discrepancy invites the use of various mechanisms for accounting for the superior performance of another. If subjects in fact do make implicit probability estimates, it is clear from the low average discrepancies in Table 3 why there is correlational evidence of social comparison in the ascending and random conditions. Second, even ignoring social comparison processes, the primacy effect could stem from an initial impression of high ability followed by an accounting for later variations in performance in terms of various nonability factors.

If something like this latter factor were the only essential factor underlying the primacy effect, then the relative superiority of the descending SP should be manifest even if the subject-perceiver were to make no attempt to solve the problems herself. If the superiority judgment has anything to do with implicit probability estimates, furthermore, subjects should at least be able to make estimates approximating the Atkinson model when asked to do so. The fourth experiment was an attempt to examine these two implications.

EXPERIMENT IV

Method

The fourth experiment was very much like the second, with a few crucial exceptions. The standard performance film was shown to four groups of subjects ranging in size from 7 to 11; the total sample was 37, though 3 subjects were discarded for not following instructions. All subjects were again female volunteers from the introductory psychology course. The design included the three basic conditions: descending success, ascending success, and random success.

As distinct from each of the preceding experiments the subjects were instructed not to solve the first series of problems themselves but instead to try to predict, prior to the display of each problem, the probability that the filmed SP was correct on the upcoming problem. These probability judgments could therefore take into account the perceived difficulty of the preceding problems and the announced success-failure feedback for the SP. They could not take into account the difficulty level of the particular problem on which probability was being predicted. In an effort to convey information about problem difficulty comparable to the preceding experiments, the subjects were also given feedback information about the subject whose back faced the camera. This was the same pattern of 10 correct responses that the subjects in earlier experiments received. In addition, a bogus norm was reported for each problem in order to emphasize that the problems were of equal difficulty. This feedback indicated that from 22% to 27% of a normative undergraduate group solved a given problem.

Having made probability estimates rather than attempting to solve problems of the first series themselves, the subjects were then shown the second portion of the film and asked to predict whether the SP was correct on each of the 30 successive problems. This was followed by a questionnaire similar to those used in the preceding experiments. This questionnaire included an intelligence rating, a recall of SP performance on the first series, and a guess of how many problems out of the second series of 30 they themselves would have correctly solved.

Results

Ability attribution and recall. As Table 4 shows, there were no significant variations in the subjects' predictions of the SP's performance on the second series. The predictions were again lowest for the ascending SP, but for the first time the random SP was expected to do better than even the descending SP.

TABLE 4

VARIATIONS IN ABILITY ATTRIBUTION: EXPERIMENT IV

Condition	n	Prediction summary	Rated intelligence[a]	Recall of SP performance[b]
Descending	10	18.11	7.70	20.60
Ascending	14	17.36	5.60	12.50
Random	10	18.40	6.70	14.60

[a] Descending versus ascending: $t = 3.09$, $p < .01$; other comparisons nonsignificant.
[b] Descending versus ascending: $t = 6.45$, $p < .001$; descending versus random: $t = 3.96$, $p < .001$.

None of these differences approached significance.

In spite of the finding of no difference in prediction scores, there once again were differences in rated intelligence. The descending SP was seen as significantly more intelligent than the ascending SP ($p < .01$), and the random SP was almost exactly intermediate between the two systematic trend conditions (and not significantly different from either).

As one considers the recall of number correct on the first series, the superiority attributed to the descending SP is striking, and the position of the ascending SP is unusually low. Even though the descending and ascending SPs both were correct on 15 out of 30 problems, the descending SP is recalled as being correct on 20.6 problems and the ascending SP on 12.5, a difference significant well beyond the .001 level, and one that is much larger than any observed in the preceding experiments. The task of recording probability-of-success estimates forced the subjects to attend carefully to each trial and its outcome. Instead of increasing the accuracy of recall, this increased attentiveness apparently led to an exaggeration of the primacy effect on the recall measure.

Probability of success estimates. Each subject made a separate judgment of probability of success on the next problem after learning how the SP had one on the immediately preceding problem. The average judgments by item (trial) and by condition are plotted in Figure 2, along with the curves expected from the simple Atkinson formula. Several features of this figure deserve comment. There is, first, the perplexing fact that the ascending and descending SPs were seen to differ even

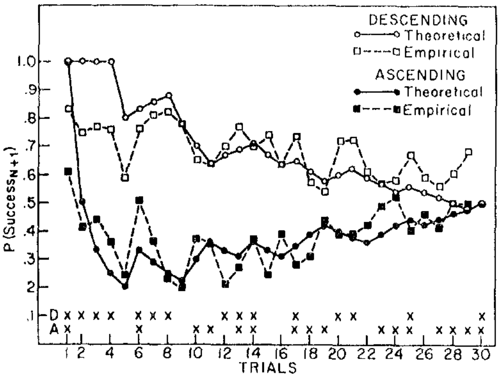

FIG. 2. Judgments of probability of success on the next trial (Experiment IV): theoretical and empirical functions.

after the first trial, on which they had both received "correct" as feedback. This is a highly significant difference and one that cannot plausibly be dismissed as "chance." The experimenter might have conveyed in some unwitting way that the ascending SP was not very intelligent. Such a possibility of "experimenter bias" would, of course, vitiate the significance of the results. A more likely (and admittedly more congenial) alternative is that the differences on the first trial are the result of differential emphasis on the range of numbers available to the subjects for making their probability estimates. The sheet on which the probability of success ratings were to be recorded contained a misleading code that emphasized the ratings of 0, 5, and 10—chances correct out of 10 tried. The experimenter recalls placing greater emphasis in the descending than in the ascending conditions on the full range of numbers from 0 to 10. Of the 14 subjects in the ascending condition, 1 chose the rating of 10 and 9 chose 5 after the first trial. Of the 10 subjects in the descending condition, 2 chose 10 and none chose 5, the rest choosing 7, 8, or 9. This difference in the salience of rating alternatives apparently lasts but a short time. As the raters have more and more information to go on after the first trial, there is an equivalent use of intermediate ratings in both ascending and descending conditions. Also, those subjects in the ascending condition using 5 as an initial rating ended up making higher probability ratings than those initially responding with 7, 9, or 10 ratings. Whatever the origin of the rating difference between ascending and descending conditions on the first trial, subsequent judgments seem unrelated to these initial estimates.

It is apparent from Figure 2 that the empirical averages map very closely the theoretical curves for the two conditions from the second to the twenty-ninth trial. This, the authors believe, is the most striking feature of the graph: that, when instructed to make probability estimates, subjects do a remarkable job integrating previous successes, and appear to be implicitly using the Atkinson formula in their judgments. The subjects are obviously sensitive to the previous history of success and failure, though they tend to "overshoot the mark" in responding to the more recent successes or failures. That is, a trend upward or downward in the theoretical curve is accompanied by an exaggerated trend upward or downward in the empirical curve.

SP prediction and own performance estimate. In Experiment III, it was observed that subjects in the descending SP condition did not, like those in the remaining conditions, perceive any necessary relationship between the score predicted for the SP and the hypothetical score they would receive if they had worked on the second series of problems. This was interpreted as evidence of comparison cessation by subjects in the descending SP condition. The observed difference between ascending and descending conditions was again observed in the present experiment, with the random condition once again in an intermediate position. The product-moment correlation coefficients are presented in Table 5. The fact that there was the same correlation difference in the preceding experiment, where all subjects at-

TABLE 5

CORRELATIONS BETWEEN ESTIMATE OF HYPOTHETICAL OWN SCORES AND PREDICTION OF SP SCORE ON SECOND SERIES: EXPERIMENT IV

Condition	n	Product-moment correlation
Descending	10	.09
Ascending	14	.67*
Random	10	.28

* $p < .01$.

tempted to solve the first series of problems, and in the present one, where subjects have no history of personal solution effort and feedback, poses difficulties for the interpretation that comparison cessation is involved in the descending SP condition. In order for this interpretation to remain valid, one must assume that the subjects in the descending condition have a fairly clear idea about what their own performance would be on the early trials even though they are making probability judgments of SP success rather than attempting to solve the problems themselves. If the subjects were to identify with the second filmed subject (back to the camera) or with the average college student (as described by the normative information given after each trial), the comparison hypothesis would still be appropriate.

While the underlying causes are by no means clear, the fact remains that in the two experiments in which subjects were asked to guess how well they would do on the second series, these estimates are linked to predictions of SP success very closely in the ascending condition, to an intermediate extent in the random condition, and not at all in the descending condition. Something about the ascending condition promoted feelings of comparability that are lacking in the descending condition.

Discussion

The results of Experiment IV replicate only partially the consistent findings of the preceding experiments. Attributions of intelligence, and especially the recall of SP performance on the first series, fall into the now familiar pattern in which the descending SP is judged superior to the ascending SP. However, this customary difference does not reveal itself on the prediction summary scores. The major difference between the fourth experiment and the preceding three is that the subjects in the fourth were given no opportunity to solve the problems. It is possible, therefore, that the prediction summary scores only spread apart when feelings of competition are involved, and that such feelings emerge only when (*a*) the SP is seen to lie within the subject's comparison range, characteristic of the ascending condition, and (*b*) the subject solves the problems along with the SP. Perhaps the arousal of competitive feelings results in a tendency to belittle the SP—to predict poor performance on the next series. When the subjects do *not* attempt to solve the first series of problems, their feelings about comparison are less clearly linked to competitive feelings. Therefore, there is less motivational pressure to belittle the SP's performance.

In general terms it is suggested that more ability is attributed to the descending SP than to the ascending SP for at least two and possibly three reasons. First, one may posit a purely cognitive factor unrelated to social comparison processes, but stemming from the special character of ability as a disposition. The authors have already speculated that decisions about ability are made quickly and often on the basis of inadequate or misleading information. The salience of early successes and failures in probability of success estimates may operate to exaggerate modest ability differences on the early trials. Once a certain level of ability is assigned, subsequent variations in performance are apt to be accounted for in terms of variations in motivation, luck, circumstances, and so on. To the extent that this speculation is true, at least some of the recall and attribution results follow naturally: early information about ability tends to carry more weight than later information.

The remaining two factors involve two processes of social comparison, a purely self-evaluative process and a self-enhancing process. Several of the studies reported in Latané (1966) suggest that social comparison can serve the motive to gain more precise information about one's own standing as well as the motive to see oneself in the best possible light. The role of social comparison is probably not independent of the fact that early information is highly salient in the ability realm. Because comparison with the gifted descending SP does not particularly satisfy the subjects' needs for precise self-evaluation, they cease to compare their ability with the SP's. This cessation of comparison reduces the normal tendency to see the SP as more similar to the self than he actually is. The judged proximity of the ascending SP, on the other hand, generates comparison pressures that cause the subject to maintain similarity. As a result, the subject confronting an ascending SP tends to recall his success as less impressive than it actually was and avoids attributing high intelligence. In addition, to the extent that the subject does predict a high level of performance on the second series, she thinks that she herself would do well.

It should again be stressed that we are talking only about informational comparison pressures—those directed toward precision of self-evaluation. It is another question whether esteem-protecting social comparison processes are also involved. The results of the last experiment suggest that they may be, but the evidence is far from compelling. In Experiment IV the subject's needs to protect her self-esteem should be minimal since she herself is not attempting to solve the problems. The results of this experiment are distinctive in showing little or no relationship between ratings of intelligence and recalled performance on the one hand, and predicted performance on the other. The subjects do not seem to "care" about how well the SP does on the second series, and the subjects in the ascending condition are less inclined to drag the SP down to their level. Although the correlational data still show a concern with comparison in the ascending condition, this may be explained in terms of the need for evaluative information rather than the need to bolster self-esteem.

Since the preceding experiment was the first in which subjects did not themselves solve the first series of problems *and* the first in which the summary predictions did not differ in the ascending versus the descending condition, it is tempting to relate these two "firsts." But the evidence involves comparison across experiments and thus leaves something to be desired. Since the ascending-descending difference is at least in the expected direction on the summary prediction index, one must certainly consider the possibility of a Type II error: a true difference on the prediction data may exist even when the subject does not solve the initial series of problems, and the results of Experiment IV simply fail to reveal this difference.

In order to put this possibility to a more precise test, a fifth experiment was conducted to permit a direct comparison between subjects who solve the first 30 problems and subjects who instead predict the trial-by-trial performance of the SP. The experiment was also conducted in such a way as to eliminate the confounding of conditions with sessions and to see if males would produce the same pattern of results as females.

EXPERIMENT V

Method

A total of 69 male and 71 female undergraduates appeared as subjects in four group-administered experimental sessions. The number of males and females was approximately equal in each session, and the four sessions ranged in sample size from 32 to 37. In two of the sessions subjects were instructed to attempt to solve the first 30 problems; in the remaining two, subjects made probability-of-SP-success estimates before each of the first 30 problems instead of attempting to solve the problems themselves. In this respect, the experiment represented a combination of Experiments II and IV.

Within each session, subjects were randomly assigned to either the ascending or descending success condition. This was accomplished by the distribution of interleafed experimental booklets that alternately contained ascending or descending feedback information referring to the SP's performance. No random feedback conditions were included in this experiment.

After the subjects had worked through the first 30 problems, either attempting to solve them or estimating the probability of SP success, identical verbal instructions were given to all groups and the experimenter showed the standard film of the SP in the act of solving the second 30 problems. As in all previous experiments, the subjects predicted success or failure for each solution attempted by the SP and were told nothing about their accuracy. At the completion of the film, the subjects were asked to return to their booklets containing the second series of problems and to record for each problem whether *they* would have been able to solve the problem or not. The subjects then made the customary ratings of intelligence and recorded their recollections of the number of items on which the SP had been correct over the first 30 trials.

As in Experiment II, subjects in the solution conditions were given a standard pattern of correct multiple-choice alternatives so that they could evaluate their own performance. This meant, of course, that the level of their performance could not be precisely controlled. The resulting scores again covered a wide range, from 2 to 20 correct. Also replicating Experiment II, there was no relationship between the number of problems "correctly" solved by the subject and the number of subsequent successes he predicted for the SP. Since this was the case, and since there were no differences across conditions in the average number solved (the condition averages ranged from 11.61 to 12.78), no attempt was made to discard subjects with atypically high or low performance scores.

Results

The condition means for the three dependent variables of primary interest appear in Table 6. Analysis of variance summaries for these same variables are in Table 7. Several things are apparent from these results. First of all, there are clear replications of the previously observed effects of performance pattern on each measure of ability attribution. Once again, relative to the ascending SP, the descending SP is recalled as having done better, is perceived as more intelligent, and is expected to do better on the second series of problems. The most striking of these effects is the distortion in recalling the number of items the SP solved correctly on the first series ($F = 43.34$, $p < .001$); difference in attributed intelligence is also highly reliable ($F = 16.67$, $p < .001$); and the prediction of success on the second series is barely significant ($F = 4.07$, $p < .05$). One may infer from the replication feature of these results that the prior findings were not the effects of subtle experimenter bias or other effects arising from the partial confounding of sessions and conditions that characterize Experiments II and IV.

Effects of sex and problem-solving experience. In general, male subjects predict a lower level of SP performance than female subjects. They do not, however, attribute less intelligence to the SP or recall her performance as being significantly lower than do the females. It is difficult to know what to make of this male pessimism. Males do not solve any fewer problems than females in the solve conditions. The difference cannot, therefore, be attributed to the use of one's own performance score as a basis for prediction. Whatever the reason for the main effect of sex on prediction summaries, subjects of both sexes agreed in attributing greater intelligence and better performance to the descending SP.

Turning to the effects of solving the first 30 problems versus estimating the probability of SP success, again there is a statistical main effect. Subjects who attempt to solve the problems themselves tend to predict lower SP performance than those in the estimation conditions. This could easily be explained as a reflection of differential familiarity with the difficulty level of the problems and will not be further considered.

In the discussion of Experiment IV, the authors speculated concerning the failure of

TABLE 6

VARIATIONS IN ABILITY ATTRIBUTION: EXPERIMENT V

Condition	n	Prediction summary	Rated intelligence	Recall of SP performance
Female				
Solve				
Descending	18	18.11	7.50	16.22
Ascending	20	16.85	6.50	13.50
Predict				
Descending	15	19.53	7.00	18.33
Ascending	18	17.83	6.67	13.00
Male				
Solve				
Descending	17	16.28	7.56	16.47
Ascending	18	14.89	6.44	11.78
Predict				
Descending	20	16.90	7.20	16.60
Ascending	14	16.71	6.21	14.00

TABLE 7

ANALYSIS OF VARIANCE SUMMARIES: EXPERIMENT V

Source	Prediction summary		Rated intelligence		Recall of SP performance	
	MS	*F*	*MS*	*F*	*MS*	*F*
Sex (A)	123.73	11.77**	.14		10.51	
Predict-Solve (B)	51.17	4.87*	1.84		33.87	2.89
Ascending-Descending (C)	44.69	4.25*	25.56	16.67***	508.15	43.34***
A × B	.00	—	.14		1.18	
A × C	4.18	—	1.27		1.26	
B × C	1.27	—	1.36		1.03	
A × B × C	5.86	—	.64		47.73	4.07*
Error	10.52		1.53		11.72	

* $p < .05$, $df = 1/133$.
** $p < .01$, $df = 1/133$.
*** $p < .001$, $df = 1/133$.

TABLE 8

CORRELATIONS BETWEEN ESTIMATES OF HYPOTHETICAL OWN SCORE AND PREDICTIONS OF SP SCORE ON SECOND SERIES: EXPERIMENT V

Condition	n	Product-moment correlation
Female		
Solve		
Descending	18	.49*
Ascending	20	.55*
Predict		
Descending	15	.38
Ascending	18	.81**
Male		
Solve		
Descending	17	.39
Ascending	18	−.11
Predict		
Descending	20	.51*
Ascending	14	.62*

* $p < .05$.
** $p < .001$.

prediction summary scores to spread apart when subjects were instructed to estimate probabilities rather than solve the problems. To what extent might this absence of a discrepancy reflect reduced competitiveness or the absence of comparison motivation? A major purpose of the fifth experiment was to permit a direct comparison of *solve* and *estimate* conditions on the summary prediction scores. Since the interaction between ascending-descending and solve-estimate does not approach significance, there is no reason to attribute the lack of an ascending-descending difference in Experiment IV to the reduction of comparison motivation. Estimate and solve conditions in the present experiment appear to contribute equally to the significant main effect of the ascending-descending variable—especially when only the female subjects are considered (see Table 6). The most reasonable conclusion is that it makes no difference, either in the prediction of performance or the attribution of intelligence, whether the subject himself tries to solve the first 30 problems or attempts to estimate the SP's success on each upcoming trial.

Prediction summary and self-estimation score. In both Experiments II and III the correlations were examined within each condition between the subject's prediction of SP performance (Series 2) and her estimate of how many of the same problems she herself would have been able to solve. In both cases high positive correlations were observed in the ascending success condition and negligible correlations in the descending condition. These differences have been treated as evidence for greater comparison motivation in the ascending condition. Correlations between the same two summary scores were computed in the present experiment and the coefficients are reported in Table 8. The female subjects show the same trend that was previously observed—higher coefficients in the ascending conditions—but the males do not. In the male-solve conditions there is even a clear reversal of direction.

Probability estimation trends. Figure 3 presents a striking replication of the probability estimation data graphed in Figure 2. Because of the fact that ascending and descending patterns were run in the same sessions, there was obviously no reason to expect the Trial 1 difference obtained in Experiment IV to replicate. The slight difference between the ascending and descending starting points does not in fact approach significance. The significant difference in Experiment IV between the Trial 1 estimates may thus be interpreted as a harmless artifact—quite probably the result of unwitting variations in instructions.

In order to facilitate meaningful comparison with Figure 2, Figure 3 presents the data for females only. The male curves were very similar in shape, but displaced downward ($F = 3.08$, $df = 1/63$, $p < .10$), in line with their pessimistic predictions of SP success on the second series of problems. The two figures

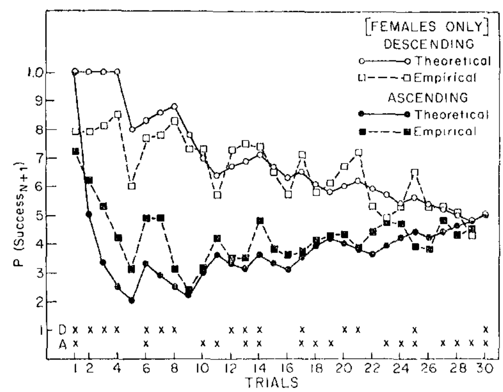

FIG. 3. Judgments of probability of success on the next trial (Experiment V): theoretical and empirical functions.

together, and the male data as well, show the subjects' remarkable sensitivity to the SP's past performance history in predicting her subsequent performance. On the average, at least, they are faithful users of the "Atkinson ratio."

Discussion of Experiments I through V

The results of the first five experiments are summarized in Table 9. We are unquestionably dealing with a rather robust and replicable finding that the performer who does better at first is seen as more capable than the performer who excels later in the series but solves the same number of problems. The last experiment established rather firmly that this finding does not depend in any important way on the arousal of competitiveness or on the operation of esteem-maintaining cognitive work. The major results tend to be the same whether or not subjects are in a position to compare their own performance with that of the SP. As far as the recall-of-SP-success measure is concerned, the difference between ascending and descending conditions is, if anything, greater when subjects are concentrating on the SP's performance and not solving the problems themselves. The major findings cannot simply reflect the subject's inattentiveness after the first few trials or his failure to monitor the belated performance spurt of the ascending SP. As the probability estimation curves show, subjects follow the Atkinson ratio very closely throughout the entire series of 30 trials. They could hardly reproduce this ratio so faithfully if they had tuned out early in the series.

The results of the fifth experiment weaken the case for a social comparison explanation since it is hard to see how subjects who do not solve the problems themselves could have a clear basis for comparing themselves with the SP. And yet, insofar as the subject has some idea about her performance level, perhaps through identification with the second subject on the film, social comparison could still be a factor influencing the results. It is hard to ignore the fact that in three separate experiments, female subjects confronting a female SP seem more intent on maintaining comparison in the ascending than in the descending condition, as measured by the

TABLE 9

SUMMARY OF MAJOR ABILITY ATTRIBUTION DATA: EXPERIMENTS I–V

Experiment	Prediction summary	Rated intelligence	Recall of SP performance
I			
D	20.2** ***	7.6[a]	17.8*
A	17.2	7.4	15.2
R	16.4	7.2	16.9
II			
D	18.7** *	7.5** *	16.8** *
A	15.8	6.4	14.7
R	16.0	6.8	14.8
III			
D	17.3[b]	8.6**	16.7*** ***
A	16.7	7.6**	14.0
R	16.9	8.4	13.7
IV			
D	18.1	7.7**	20.6*** ***
A	17.4	5.6	12.5
R	18.4	6.7	14.6
V[c]			
D	17.6*	7.3***	16.8***
A	16.6	6.5	13.0

Note.—Abbreviations: D = descending, A = ascending, R = random.

[a] The perceived discrepancy between rated SP and rated own intelligence is significantly greater in the ascending than in the other conditions.

[b] The perceived discrepancy between sum of item-by-item predictions for SP and for self is greater in the descending than in the ascending conditions.

[c] Combining both sexes and ignoring the estimate-solve variation.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

correlations between predictions of SP success and estimates of own hypothetical success. This perception of greater proximity to the ascending SP may in some way reflect the implicit probability of success estimates made by each subject. As Table 3 showed, the subject who solves 10 problems randomly distributed throughout the first series of 30 would end up with a probability-of-own-success curve that was much closer to the ascending curve than to the descending one. The nonsolving subject who identifies with the second female on the film would be in a similar position. Thus it is possible to see how a concern with predicting trial-by-trial success could produce data for the subject inviting social comparison with the ascending but not the descending SP.

Even if one takes the correlational data seriously, however, social comparison may be the *result* of a cognitive judgment that the ascending SP is similar to the subject in ability rather than a process causing or affecting such a judgment. The fact that subjects place greater weight on the first few clues about ability than on subsequent information could stem directly from the probability-of-success estimates. It is in the nature of such estimates, according to the Atkinson ratio, that early successes are more heavily weighted than later ones. That is, early successes and failures produce more abrupt changes in the prediction ratio.

The Atkinson ratio refers to performance prediction and is relevant to decisions about ability. It is not clear at this point whether the present experimental findings are peculiarly characteristic of ability attribution. Might not such primacy effects also appear when information concerning other personal attributes is presented in the same sequential patterns? In the introductory section, it was suggested that ability *may* be somewhat unique because unlike other personal dispositions abilities are judged to be durable tenants of the personality even though their presence may be erratically manifested. Because of their relative stability, as dispositions go, abilities may serve as important anchor points in the construction of an impression of a person. Perhaps it is not surprising that we overstress early evidence that bears on these anchor points so we can move on and make other contingent inferences about motivation, adjustment, and other personal traits.

Such a line of reasoning helps to explain why talent is readily inferred from relevant performance data that are revealed early in the impression-formation process—before the perceiver has already developed the "anchoring" hypothesis that talent is absent. As many students of cognitive development have pointed out, it is easier to form an impression than to change one. This may be especially true of stable, anchoring dispositions like ability.[4]

EXPERIMENT VI

The evidence thus far suggests that early information about ability is overweighted and leads to premature and persistent attributions. But this should only be the case when the perceiver must base his judgments solely, or almost entirely, on the performance of the SP in the restricted situation that has been devised. The subjective probability formula assumes complete ignorance at the outset and proceeds from there to deal only with the cumulative pattern of successes and failures. One would not expect such a primacy effect if the subject were predicting his own subsequent performance on the basis of a prior history of descending, ascending, or random performance. One may assume that the average college student subject has rather well-anchored notions about his own relative intellectual ability. When confronted with a series of items from a test "designed to discriminate at the highest levels of intelligence," the subject should be more concerned with evaluating the difficulty of the test than with reevaluating his own intelligence. This suggests a probable difference in responding to different patterns of success-failure feedback about own performance. The descending success subject learns that his early control over the task was something of an illusion. In spite of continuing to try hard, he has moved into a phase where success eludes him; he is likely to conclude that the test is harder than he thought and be rather pessimistic about future successes. The ascending success subject, on the other hand, seems to have gained control over the task. After a rocky beginning he has "caught on," as it were, and should end the first series with the general appraisal that the items are less difficult than they first appeared. The random success subject has never had control over the task and can only expect to do about the same—on some combined basis of ability and luck—on the following series.

Additional factors may also be involved. Quite aside from his judgments of item difficulty, the ascending success subject may feel he has a "lucky streak" going and simply

[4] Others (Anderson, 1965; Asch, 1946; Luchins, 1957) have observed primacy effects in person perception and perhaps it is a more general phenomenon than this discussion implies. However, the demonstrations of primacy have thus far involved verbal ratings based on verbal stimulus materials—especially the serial adjectives of Anderson and Asch—and it is not clear whether the processes underlying primacy in these settings are the same as those in the more involving behavioral setting of the present experiment.

continue, in the prediction series, to "bet on himself." In contrast, the descending success subject may be more inclined to husband his chips and take a self-protective, pessimistic view of the immediate future.

On either or both of these grounds, one should hypothesize that predictions of success on the second series would be highest among ascending subjects, next highest among random subjects, and lowest among descending subjects. However, the ascending and descending subjects do have something in common that the random subjects do not share: at one point they have controlled the task to an extent that cannot easily be accounted for by luck. The descending subject may harbor the hope, therefore, of regaining control at some point on the second task. This possibility makes it difficult to predict the relative position of descending versus random subjects. In any event, Experiment VI was designed to test the hypothesis that the ascending success subject would predict greater success for himself on a second problems series than either descending or random success subjects would, in contrast to the results of the preceding five experiments. No clear predictions were made concerning the relative positions of subjects in the descending and random conditions.

Method

The experiment, following closely the procedural steps of Experiment I, involved a subject and an experimental accomplice working in parallel on the 30 items of Series 1. Each received feedback, as in the earlier experiment, but the patterns were reversed so that the accomplice received a random pattern of 10 successes and the subject received either a descending, ascending, or random pattern of 15 successes. After the subject and the accomplice had completed this first series, the experimenter (a female undergraduate) explained that she was really interested in how well each could predict the other's performance on a second series of similar items.

The second series of items was contained in a pamphlet, copies of which were placed face down in front of the subject and the accomplice. Each was to look at each item for 20 seconds and record at the bottom of the page the number of the answer he thought to be correct plus his prediction as to whether the partner would be correct or not on that problem. Each was told that he would receive $.15 for each correct answer and for each correct prediction, but that he would lose $.05 each time he answered or predicted incorrectly. Both were assured that no money would actually be taken from them and told that the greatest sum won by a subject thus far was $4. When the series was complete, the experimenter asked the subject and the accomplice to flip through their problem booklets and to indicate beneath each question whether they thought their own answer was correct or not. Once again, they were promised $.15 for correct predictions and would be docked $.05 for incorrect ones.

The accomplice was then led to a separate room and the subject completed a postexperimental questionnaire. The subject was finally interviewed by the experimenter who probed for suspicions concerning the true nature of the experiment or for doubts that the feedback had been valid. No suspicions or doubts were elicited. The subject was then debriefed and given $1 with the apology that no true amount of winnings could actually be assigned.

Subjects and accomplices. The subjects were male introductory psychology students who participated, as did those in the preceding experiments, to help fulfill an experimental requirement for the course. One of these subjects failed to follow experimental instructions. The remaining 30 were assigned equally to the three experimental conditions. Either of two undergraduate accomplices confronted each subject, the assignment depending partly on their respective daily schedules, but also on whether one accomplice might have known the subject. Accomplice B appeared with 16 subjects (6, 6, 4 by conditions) and Accomplice R appeared with 14 (4, 4, 6). No reliable effects were associated with the particular accomplice, so they will hereafter be treated as interchangeable in discussing the data.

Results

Predictions of success on Series 2. It was predicted with considerable confidence that the ascending success subject would expect to do better on the second series than either the descending or random subject. As Table 10 shows, this prediction was confirmed ($F = 10.89$, $df = 1/27$, $p < .01$). Conflicting reasons made it difficult to predict whether the descending or the random subject would have more optimism concerning his successes on the second series. The results show that

TABLE 10

PREDICTION OF NUMBER CORRECT ON SECOND SERIES: EXPERIMENT VI

Condition	n	Own performance[a]	Partner's performance[b]
Descending	10	17.2	14.2
Ascending	10	20.4	13.4
Random	10	15.1	12.2

[a] Analysis of variance across conditions: $F = 15.16$, $df = 2/27$, $p < 001$. Ascending versus descending: $F = 10.89$, $df = 1/27$, $p < .01$; ascending versus random: $F = 29.88$, $df = 1/27$, $p < .001$; descending versus random: $F = 3.87$, $df = 1/27$, $p < .10$.

[b] None of these comparisons is significant.

TABLE 11

RECALL OF NUMBER CORRECT ON FIRST SERIES BY BLOCKS OF 10 ITEMS: EXPERIMENT VI

Own score	Trial block 1–10 Own	1–10 Partner	11–20 Own	11–20 Partner	21–30 Own	21–30 Partner	Total Own	Total Partner
Descending								
Recalled	6.8	3.5	5.0	3.3	4.0	3.5	15.8	10.3
Actual	7	4	5	2	3	4	15	10
Ascending								
Recalled	3.2	3.2	5.1	3.1	5.9	3.3	14.2	9.6
Actual	3	4	5	2	7	4	15	10
Random								
Recalled	5.0	3.8	4.5	2.9	4.8	3.2	14.3	9.9
Actual	5	4	5	2	5	4	15	10

Note.—n = 10 in all conditions.

the descending SP is somewhat more optimistic, though the difference falls slightly short of significance ($F = 3.87$, $df = 1/27$, $p < .10$). Table 10 also shows that the subject's pattern of feedback on Series 1 does not affect his predictions of his partner's success on Series 2. The small differences in the final column are not significant.

Recall of performance on Series 1. From Table 11, one may conclude that (*a*) the subjects quite accurately recall the pattern of their own successes and failures on the first series; (*b*) there is no significant tendency to distort recall of the overall level of one's own performance, though the average descending subject recalls slightly more than the actual number of successes whereas the remaining subjects recall slightly less; (*c*) a subject's own performance does not affect his generally accurate recall of his partner's performance. It is quite intriguing to note that even though the ascending subject errs in underrecalling his successes on Series 1, he predicts high performance on Series 2. The average descending subject thinks he will get 1.4 more items correct on the second series as compared with a figure of 6.2 for the average ascending subject. The expectation of doing well does not grow out of distorted recall of past successes as it apparently does when the focus is on the performance of a descending or ascending SP.

Ratings of intelligence. If the reasoning that led up to this experiment is correct, subjects should not reevaluate their intelligence as a function of their performance pattern on the first series. In accord with this expectation they do not do so. On an 11-point scale ranging from "well below average Duke student" (in intelligence) to "well above average Duke student," the mean self-ratings were descending 4.95, ascending 5.65, and random 6.30. These variations are not significant and the ordering of means is quite different from the ordering of prediction scores. Ratings of the partner's intelligence were slightly lower than the self-ratings in each condition, and again differences among these means did not approach significance.

Ratings of motivation and of item difficulty. On the final questionnaire subjects were asked to rate their "concern with doing well" and the extent of their "concentration in Series 1," both on 11-point scales. Since ratings on these two scales turned out to be comparable, they were combined into a single index of perceived motivation. Although the overall F value for the three conditions was not significant ($F = 1.95$, $df = 2/27$), the ascending subjects acknowledged significantly higher motivation than did the descending subjects ($t = 2.68$, $df = 19$, $p < .02$). Subjects in the random condition rated themselves almost as low in motivation as did the descending subjects (13.5 versus 13.1), but their ratings were much more variable and the comparison with the higher (15.2) ascending subject mean did not approach significance. The observed differences in attributed motivation appear to be a logical consequence of the instructions and the pattern of events. The descending subject can account for the decline in his performance by emphasizing that he lost interest or could not concentrate, thus protecting his stable view of his intelligence. It is also possible that he did in fact lose interest when his performance began to deteriorate. The ascending subject may in fact have tried harder, after the initial evidence of poor performance, and the subsequent success was to him his reward.

An additional way of accounting for the pattern of successes and failures would be to perceive "justifying" patterns of item ease or difficulty. In all the experiments reported, including the present one, a stressed feature of the instructions was that the items were chosen to be equal in difficulty. Also, in those experiments where the subjects themselves received feedback, the pattern was random. As a check on their agreement with this claim, subjects were asked to indicate whether they thought the early items were harder than the later ones (on the first series), whether they were equal in difficulty throughout, or whether they thought the later items were harder. In the preceding experiments there was a slight (nonsignificant) but recurrent tendency for subjects in the ascending SP conditions to see the later items as easier than in the other conditions. Somewhat surprisingly, subjects in the descending SP condition almost never indicated that they thought the early items were easier. In the present experiment there were tendencies in both directions: six ascending subjects thought the later items were easier, only one thought they were harder; six of the descending subjects thought the early items were easier, none thought they were harder; seven of the random subjects thought the items were equal in difficulty throughout. There appears to be, then, more contingency between performance pattern and judgments of item difficulty when the judge's own performance is involved than when he is focusing on someone else's. The findings on both the motivation measure and the measure of judged item difficulty support the earlier assumption that subjects have well-anchored conceptions of their own intelligence and will account for variations in performance in other ways. The ascending subjects are most likely to predict a high level of performance on the second series because (*a*) they have the feeling that when they make the effort they can control this kind of task and (*b*) the items will most likely continue to be similar in difficulty level to those to which they were most recently exposed.

Discussion

By themselves the results of Experiment VI would seem to fit cozily with a number of common sense assumptions and to have little surprise value. They are very similar to those obtained by Cope (1967) and Cope and Sigall (1967) in a situation featuring competition between two persons. In both of these studies, improving performance resulted in a higher level of aspiration than sporadic performance yielding a higher average score.

The results take on greater importance in the context of the preceding experiments, helping to clarify some issues and raising a number of additional ones. First of all, Experiment VI should lay to rest any lingering suspicions that the primacy effect in the first four experiments is an artifact of the particular problems solved in the different feedback conditions. If there were some peculiar relationship between problem content and feedback pattern, one would hardly expect a reversal of results—primacy turning into recency—when the subject is predicting his own subsequent performance.

Second, one learns that the effects of performance pattern are not simply a function of some primitive prior entry effect. The order of information received obviously plays a different role depending on whether the information is about oneself or someone else. Nor does one seem to be dealing with a reversible figure illusion of memory, with primacy flipping over into recency with the shift in perspective from target person to self. In fact, the recall measure (see Table 11) still shows a slight primacy effect: the descending subject recalls doing better than the ascending subject—even though the latter predicts a much higher level of subsequent performance. By inference, predicition of another's performance primarily reflects judgments of her intelligence; when one's own performance is involved, there is greater tendency to relate performance fluctuations to variations in item difficulty. The difference is reminiscent of Heider's (1958, p. 157) proposition that perceivers are overly ready to attribute the reactions of another person to the other's personal characteristics (e.g., ability); one's own reactions, on the other hand, are attributed to the object world (e.g., the difficulty level of problems).

While the main concern has been with the theoretical significance of this research, there are some practical implications that bear mentioning. If one imagines a situation in which A is observed by B to show systematic improvement in performance, the present results imply that A will feel very confident about his own performance on future occasions, whereas B's estimates about A's subsequent performance will be relatively low. The obverse is also true: A will not be optimistic about his future performance when his immediately prior efforts have led to decreasing success; B, on the other hand, will have high expectations for A under these circumstances. Such conflicting expectations could obviously create interpersonal difficulties and misunderstandings. The improving player may wonder why he was kicked off the team precisely when he sees himself at the peak of his game. The worker whose performance is deteriorating may be traumatized by a promotion to an even more responsible and demanding position. Other comparable consequences of discrepancies in ability attribution readily suggest themselves.

Even disregarding the results of Experiment VI, the consistent evidence for a primacy effect in attributing ability to others has important implications for the recognition and reward of talent. Graduate students selected for their high GRE scores, who subsequently do poorly in graduate school, may be retained in spite of this poor performance because it is attributed to temporary motivational or neurotic factors. The "late blooming" worker or student may never get the recognition for his ability that he deserves. Others may tend to treat his later successes as instances of high motivation ("he's a plodder") or even luck.

The authors hasten to add that these extrapolations go well beyond the data available, and are offered as suggested lines for further research. These experiments have dealt with a particular series of intelligence-test problems embedded in a specific, restricted situation. There are obviously different kinds of tasks that might generate very different results. The same feedback patterns were used throughout, and the relatively poor performance of the subject relative to the SP was a constant feature in the first three experiments and, by implication, the fourth. An obvious extension would be to investigate the role of different patterns of SP performance when its overall level is lower than that of the judge's. Another extension might be to explore prediction settings in which the performance of judge and SP are contingent on each other. We believe that the primacy-recency phenomena we have observed will hold in a wide variety of performance situations, provided that: (*a*) the performance tasks are discrete and factors of learning are therefore minimized, (*b*) the difficulty level of the tasks remains roughly constant throughout, and (*c*) the tasks appear to be measures of some basic ability. It remains to be seen whether this belief is valid.

REFERENCES

ANDERSON, N. H. Primacy effects in personality impression formation using a generalized order effect paradigm. *Journal of Personality and Social Psychology,* 1965, 2, 1–9.

ASCH, S. E. Forming impressions of personality. *Journal of Abnormal and Social Psychology,* 1946, 41, 258–290.

ATKINSON, J. W. *An introduction to motivation.* New York: Van Nostrand, 1964.

COPE, V. M. The effects of sequences of winning and losing on level of aspiration in two-person competitive situations. Unpublished doctoral dissertation, University of Texas, 1967.

COPE, V. M., & SIGALL, H. Influences of winning or losing in competitive situations on level of aspiration. *Psychonomic Science,* 1967, 8, 47–48.

FESTINGER, L. A theory of social comparison processes. *Human Relations,* 1954, 7, 117–140.

HAKMILLER, K. L. Threat as a determinant of downward comparison. *Journal of Experimental Social Psychology,* 1966, Supplement 1, 32–40.

HEIDER, F. *The psychology of interpersonal relations.* New York: Wiley, 1958.

JONES, E. E., & DAVIS, K. E. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology.* Vol. 2. New York: Academic Press, 1965.

KELLEY, H. H. Attribution theory in social psychology. *Nebraska Symposium on Motivation,* 1967, 14, 192–241.

LATANÉ, B. Studies in social comparison. *Journal of Experimental Social Psychology,* 1966, Supplement 1.

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

WALSTER, E., & PRESTHOLDT, P. The effect of misjudging another: Over-compensation or dissonance reduction? *Journal of Experimental Social Psychology,* 1966, 2, 85–97.

WALSTER, E., WALSTER, B., ABRAHAMS, D., & BROWN, Z. The effect on liking of underrating or overrating another. *Journal of Experimental Social Psychology,* 1966, 2, 70–84.

(Received November 30, 1967)