

1-1-1972

## Limitations on the modeling of self-reinforcement : a social comparison analysis.

William A. DeLamarter  
*University of Massachusetts Amherst*

Follow this and additional works at: [https://scholarworks.umass.edu/dissertations\\_1](https://scholarworks.umass.edu/dissertations_1)

---

### Recommended Citation

DeLamarter, William A., "Limitations on the modeling of self-reinforcement : a social comparison analysis." (1972). *Doctoral Dissertations 1896 - February 2014*. 1441.  
<https://doi.org/10.7275/0b0y-by05> [https://scholarworks.umass.edu/dissertations\\_1/1441](https://scholarworks.umass.edu/dissertations_1/1441)

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).

UMASS/AMHERST



312066013572680

LIMITATIONS ON THE MODELING OF SELF-REINFORCEMENT:  
A SOCIAL COMPARISON ANALYSIS

A Dissertation Presented

By

William A. De Lamarter

Submitted to the Graduate School of the  
University of Massachusetts in  
partial fulfillment of the requirements for the degree of  
DOCTOR OF PHILOSOPHY

September 1972

Major Subject: Psychology

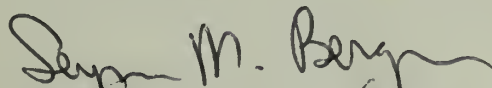
LIMITATIONS ON THE MODELING OF SELF-REINFORCEMENT:  
A SOCIAL COMPARISON ANALYSIS

A Dissertation

By

William A. DeLamarter

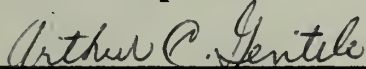
Approved as to content and style:



(Chairman of Committee)



(Head of Department)



(Member)



(Member)



(Member)

## ACKNOWLEDGMENTS

I would like to express my very sincere thanks to Dr. Seymour M. Berger for his help and advice throughout the preparation of this manuscript. His thoughtful comments and pointed criticisms have been invaluable in helping me clarify the theoretical issues which resulted in this study. Thanks are also extended to Drs. Arnold D. Well and Stephen R. Reisman for their willing and patient help throughout the preparation of this study. Special thanks are extended to Dean Arthur C. Gentile who willingly served on a committee concerned with research which was not in his area. His interest and willingness to be exposed to ideas developed in this study have been greatly appreciated. Thanks are also extended to Ms. Linda Crinklaw for serving as the model and to Mr. Robert A. Sacks for designing and constructing the experimental apparatus. Finally, very deep thanks are extended to Sheryl E. King who patiently read and commented on numerous early drafts of this dissertation and who has managed to maintain my spirits through some rather difficult times in the preparation of this manuscript.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS . . . . .	iii
LIST OF TABLES . . . . .	v
ABSTRACT . . . . .	vi
INTRODUCTION . . . . .	1
Problem . . . . .	1
Self-reinforcement . . . . .	2
Theoretical issues . . . . .	3
Previous research . . . . .	7
Social Comparison theory . . . . .	9
Modeling of self-reinforcement . . . . .	12
Proposed research and hypotheses . . . . .	20
METHOD . . . . .	22
Subjects . . . . .	22
Procedure . . . . .	22
Design . . . . .	22
Apparatus . . . . .	23
Experimental sequence . . . . .	23
Questionnaire . . . . .	26
Statistical analyses . . . . .	29
RESULTS . . . . .	29
Self-reinforcement data . . . . .	29
Questionnaire data . . . . .	33
DISCUSSION . . . . .	37
REFERENCES . . . . .	50
APPENDIX A . . . . .	52

## LIST OF TABLES

	Page
Table 1 - Summary of analysis of variance for absolute deviations from the model's standard . . . . .	31
Table 2 - Cell means by conditions for self-reinforcement data . . . . .	32
Table 3 - Analyses of variance for questionnaire data .	34
Table 4 - Cell means for helpfulness of model data . . .	35

## ABSTRACT

The present study was an attempt to apply Festinger's (1955) theory of Social Comparison to the modeling of amount of self-reinforcement. The three major independent variables in the study were model similarity (similar versus dissimilar), type of task instructions (patterned versus random order), and correct trials in a 2 x 2 x 10 repeated measures design. Subjects were asked to make predictions concerning which of two lights on a panel before them would be illuminated. Subjects were informed that the illumination of the lights would follow either a patterned or random order. Subjects were also told that they would view another subject (a model) who would demonstrate the apparatus via video tape. The model was described to the subject as either similar or dissimilar to the subject in attitude concerning the importance of doing well on the task. On the basis of Social Comparison theory, it was predicted that there would be more modeling of the similar model than the dissimilar model when the task instructions stated a patterned order to the illumination of the lights. Also, it was predicted that there would be no differences in modeling between the similar and dissimilar model groups when the task instructions stated a random order to the illumination of the lights. Subjects performed a series of twenty trials.



On half of the trials, subjects received feedback that they were correct. Data were analyzed for sixty subjects with fifteen subjects in each cell. Only the main effect of trials was found to be significant. Reasons for the failure of this study to yield significant results and suggestions for future research are presented.

## INTRODUCTION

### Problem

This study, to be described below, attempted to specify some situations which limit the modeling of amount of self-reinforcement. There is abundant evidence to be reviewed below that subjects will adopt a standard for self-reinforcement displayed by a model. However, there is very little known about those situations where the subject will ignore the self-reinforcement standard of the model. Bandura (1971) has pointed out that there are situations where another person is inappropriate as a source of cues for determining whether or not a performance is deserving of self-reinforcement. For example, it is not likely that a weekend golfer would adopt the standards of an Arnold Palmer in determining whether or not his score was deserving of self-praise. Likewise, if there were some objective, non-social cues for determining task success it is unlikely that one would adopt the self-reinforcement standards of another person. The present study, utilizing predictions based on Festinger's (1955) Social Comparison theory, attempted to delineate situations in which the subject either adopted or ignored the standards established by a model. However, before developing this line of reasoning, it is necessary to

review both the theoretical issues and the current research concerning self-reinforcement.

### Self-reinforcement

To date there has been little or no attempt to empirically investigate the uses of self-reinforcement. On the applied level, Bandura (1971) has noted that there may be a relationship between self-reinforcement and psychopathology. According to his view, deviant behaviors may be maintained through the use of self-reinforcement. Thus, many behaviors which appear to be resistant to both the traditional psychoanalytic and current behavior modification therapies may be resistant because they are maintained by the patient through the judicious use of some form of self-reinforcement. On a more theoretical level, self-reinforcement may help to maintain behaviors undergoing experimental extinction. For example, it is known that partial reinforcement schedules are more resistant to extinction than are continuous reinforcement schedules. It is conceivable that an individual being shaped on a partial reinforcement schedule provides himself with some type of self-reinforcement, regardless of its nature, on those trials where no direct reinforcement is given. Then, during extinction, the experimental subject can use self-reinforcement to maintain his behavior. The greater efficacy of partial reinforcement schedules over continuous reinforcement can be explained, using this analysis,

by arguing that partial reinforcement schedules force the subject to learn some type of self-reinforcement response during acquisition while the continuous reinforcement schedules do not. Obviously, as will be noted shortly, such a position is an anathema to researchers trained in the operant tradition. However, such a position is at least plausible. Regardless of whether or not self-reinforcement can be applied to the situations mentioned above, it is important to decide whether or not the phenomenon of self-reinforcement is useful and moreover, one must decide whether or not self-reinforcement, if it is useful, is really reinforcement.

Theoretical issues. Marston (1965, p. 255) has defined self-reinforcement as ". . . the administration of reinforcement by the subject to himself, without external controls." Bandura (1971) has argued that self-reinforcement is a theoretical concept and deserves research attention. He notes, as mentioned earlier, the use of self-reinforcement in the maintenance of deviant behaviors. In addition, Bandura argues that self-reinforcement is the process by which individuals control their own behaviors in the absence of direct reinforcement. Gewirtz (1971), on the other hand, doubts whether the term self-reinforcement is useful and whether or not self-reinforcement exists as a phenomenon. According to his view, self-reinforcement is nothing more than the maintenance of a behavior through the use of

intermittent direct reinforcement. Gewirtz bases his arguments on Bandura's statement that self-reinforcement is only relatively independent of extrinsic reinforcement. Gewirtz's operant analysis argues that because self-reinforcement is only relatively independent of extrinsic reinforcement the behaviors that self-reinforcement maintain are, in fact, maintained by this intermittent direct reinforcement. Thus, Gewirtz argues for two reasons that Bandura's conception of self-reinforcement is unnecessary. Firstly, research on self-reinforcement lacks sufficient controls to account for previous conditioning history, and therefore, evidence for the phenomenon of self-reinforcement as a unique phenomenon is lacking. Secondly, the use of terms such as self-reinforcement is not parsimonious. This latter argument states that intermittent direct reinforcement can account for the behaviors maintained supposedly by self-reinforcement, and that the concept of intermittent direct reinforcement has some basis in operant theory while self-reinforcement has no such theoretical basis. At this point it is impossible to accept one position over the other on the basis of the data. Both approaches make the same predictions. Moreover, there appears to be no feasible empirical test pitting the Bandura and Gewirtz approaches against one another. However, an even deeper issue remains unresolved. Is self-reinforcement really reinforcement?

Hilgard and Bower (1966) note that reinforcement is commonly considered to be any event which increases the



probability of a given response. However, in the research on self-reinforcement there is generally no empirical test of whether or not the use of self-reinforcement will alter the probability of a response. For example, Kanfer and Marston (1963) have investigated the use of self-reinforcement in a verbal discrimination task. In this study there is no empirical test of whether or not self-reinforcement alters the probability of any response since self-reinforcement itself is the response under consideration. In the research on the modeling of self-reinforcement (Bandura and Kupers, 1964; Bandura and Whalen, 1966; Liebert and Ora, 1968; Mischel and Liebert, 1966) the experimenters are interested in whether or not a subject can acquire a standard for self-reinforcement by observing a model. In this research there is no modification in the subject's behavior as a function of whether or not he used some type of self-reinforcement since the acquisition of a standard for self-reinforcement was the dependent measure. In addition, the subject's scores for his performance were controlled by the experimenter. Such evidence would argue that self-reinforcement does not qualify as reinforcement. On the other hand, there is some evidence that self-reinforcement, while not altering the probability of a response, can maintain a learned response at the learned level. Marston and Kanfer (1963) have utilized a verbal discrimination task where subjects were asked to select one

of four nonsense syllables. If the subject selected the correct syllable, he was reinforced. After a series of acquisition trials, the subject either continued with direct reinforcement (continued acquisition), was given control over his reinforcement (self-reinforcement), or was given no reinforcement (extinction). It was found that there was increased learning for the continued acquisition group, a reduction in correct responding for the extinction group, and neither a reduction nor increased learning for the self-reinforcement group. It is important to note that the self-reinforcement group was able to maintain the previous level of learning. The Marston and Kanfer (1963) study would appear to indicate that self-reinforcement does qualify as a reinforcement as defined earlier.

Part of the problem of whether or not self-reinforcement qualifies as reinforcement lies in the fact that researchers have confused reinforcement with a class of events labeled reinforcement. For example, various events labeled reinforcement such as lights (Marston, 1961; Marston and Kanfer, 1963; Kanfer and Marston, 1963), the word "good" (Marston, 1963), candy (Bandura and Kupers, 1964; Bandura and Whalen, 1966; Mischel and Liebert, 1966), chips to be exchanged for prizes (Marston, 1961; Bandura and Perloff, 1968; Liebert and Ora, 1968), and a rating scale (Marston, 1961) have been used as reinforcements under the subject's own control. Unfortunately, none of the above experimental designs test whether

or not these supposed reinforcers actually modify the probabilities of any response. Without empirical tests, it is difficult to determine whether or not these events are, in fact, reinforcements. Future research should be conducted which will test the effectiveness of self-reinforcement on behaviors such as those described by Bandura (1971) earlier in this paper.

Previous research. At this point, it is relevant to review the previous research on self-reinforcement. To date there have been two approaches to the study of self-reinforcement; one concerned with the effect of self-reinforcement after it has been acquired through direct training, and the other concerned with the modeling of standards for self-reinforcement. While the latter approach is of more direct relevance to the research to be presented, it is important to briefly review the former.

In the Kanfer and Marston (1963) study mentioned earlier, accuracy of the use of self-reinforcement was investigated. Subjects performed a verbal discrimination task under one of three incentive conditions and one of three reinforcement procedures. It was found that continued direct reinforcement led to increased learning of the reinforced response, self-reinforcement led to the maintenance of the previously learned response, and extinction led to a decrement in the learned response. While the high incentive condition group most accurately administered self-



reinforcement to the previously rewarded responses, incentive itself failed to affect the frequency of self-reinforcement. Kanfer and Marston (1963) investigated different variables influencing the accuracy and frequency of self-reinforcement. Using the verbal discrimination task described earlier, the experimenters manipulated (1) the stage of prior learning before introducing self-reinforcement, (2) the stringency of the instructions concerning the administration of self-reinforcement, and (3) the similarity of the experimental stimuli in the acquisition and test phases of the experiment. It was found that (1) the greater the stage of prior learning, the more accurate subjects were in the administration of self-reinforcement, (2) the greater the stringency concerning the administration of self-reinforcement, the less frequently self-reinforcement was used, and (3) changing the experimental stimuli from the acquisition to the test phase resulted in a reduction in the number of previously rewarded responses.

In another study concerned with the use of self-reinforcement, Bandura and Perloff (1968) found that self-reinforcement can maintain an effortful response. In this study, children turned a crank a number of turns in order to earn points. The subjects then selected a criterion number of points as a cut-off for self-reinforcement. A second group of subjects had the criterion of the first group imposed on them. The results of the study indicated that both

self- and externally controlled reinforcement served to maintain the cranking response.

Thus, one is led to the second type of self-reinforcement research, that concerned with the modeling of self-reinforcement. As mentioned in the beginning of this paper, the research on the modeling of self-reinforcement, be it rate, standard, or amount, leaves one with the impression that one always adopts the self-reinforcement behavior of a model. Obviously, as the example of the golfer indicated, there are times when an individual uses cues other than the behavior of a model in order to determine his own self-reinforcing behavior. It was pointed out that Festinger's (1955) theory of Social Comparison might be useful in helping to specify when the model's self-reinforcing behavior is adopted and when it is rejected. In order to demonstrate the feasibility of the Social Comparison formulation it is necessary to first review the postulates of Social Comparison theory and then to demonstrate its relevance to the current research on the modeling of self-reinforcement.

### Social Comparison Theory

Festinger (1955) has postulated a need for self-evaluation by which an individual compares himself with another in terms of ability and/or opinions. Festinger notes that when objective, non-social cues for determining the correctness of one's abilities and/or opinions are

absent one compares himself with others. The comparison other is appropriate to the extent that the abilities and/or opinions of the comparison other are similar to those of the individual. The less the similarity in abilities and/or opinions, the less likely the other will be deemed an appropriate comparison person. There are two basic independent variables in an experimental test of the Festinger formulation. They are the degree to which objective, non-social cues for comparison are present and the similarity in ability and/or opinion between the individual and the comparison other. As will be shown, the research on the modeling of self-reinforcement standards often contains no objective, non-social cue by which the subject can determine whether or not his performance is deserving of self-reinforcement. In addition, the model can generally be assumed to be similar in ability and/or opinion. Thus, there is ample reason to believe that the Social Comparison theory predictions are relevant to the self-reinforcement literature.

Bandura (1971) states that a self-reinforcing event involves social comparison processes. He notes that there are four subsidiary processes in a self-reinforcing event. The first process involves the establishment of a standard by which the individual evaluates the adequacy of his performance. The second involves a comparison by the individual with appropriate others to determine the meaning of his performance in relation to others. The third process involves

the knowledge that the self-reinforcements are under the individual's control, and the fourth process is the administration of those reinforcers to himself. In terms of the social comparison process, there appear to be four steps in the self-reinforcement sequence. The first is the decision of whether or not to perform. It is highly unlikely that a violin student will attempt to perform the same difficult composition performed by the master. In this situation, the behavior which might be potentially self-reinforced is not even attempted. Assuming that the student decides to perform the task, the second step concerns the performance itself. Is the student's performance like that of the violin master? In the third step, the individual must decide if his performance is deserving of self-reinforcement. If the student decides that his performance is equal to that of the master, he may deem the performance worthy of self-reinforcement and then administer it to himself. At this point the self-reinforcement sequence is complete. However, it is possible to reverse the roles of the student and the teacher in the above example. While it is probable that the master will be able to match the performance of the student, it is unlikely that the master will administer self-reinforcement for such a match. In this case the behavior probably is not deemed worthy of self-reinforcement. Nevertheless, it does point out that self-reinforcement may be neither deserved nor administered if the task is perceived as too simple.



The final decision of whether or not to administer self-reinforcement is of interest to most researchers. The decision of whether or not to perform is generally excluded by the experimental situation and the task is generally designed so that the performance of the subject matches that of the model. As will be noted shortly, the decision of whether or not the performance is deserving of self-reinforcement may be expanded to include the question of how much self-reinforcement is deserved.

### Modeling of self-reinforcement

To date there have been two basic types of studies concerned with the modeling of standards for self-reinforcement. The first, exemplified by Marston (1965), is concerned with the adoption of the model's rate of self-reinforcement. Subjects were asked to respond with the word "good" as a verbal reinforcer whenever they felt that they had responded with the most popular word association to a stimulus word. After measuring a base rate of self-reinforcement, subjects were exposed to a model via audio tape. High versus low rate of self-reinforcement of the model was varied by having the model either increase his rate of self-reinforcement over trial blocks (high rate) or maintain a constant rate of self-reinforcement over trial blocks (low rate). In addition, there was a control group which did not hear a model. The situation was a co-learning

task where the subject and the model alternated giving responses. The degree to which the subject overtly responded was also varied. In one condition, the subjects responded with both the word association and self-reinforcement. In the second condition, subjects did not verbalize any self-reinforcement while in a third group subjects gave neither word associations nor self-reinforcement. The results indicated that the high and low rate groups differed from the control group even though they did not differ from one another. This finding is indicative of a general modeling influence. The degree of overt responding failed to yield a main effect, indicating that overt responding is not necessary to obtain imitation of the model's rate of self-reinforcement. In terms of the Social Comparison analysis, one would expect to find the model having some influence since there were no objective, non-social cues for determining task success in this study. The only way in which the subject could evaluate his own performance was in terms of the model's performance. Since there was no reason for the subject to infer any differences in ability and/or opinion to perform the task, the subject would probably feel that his responses were as correct as the model's and thus deserving of reward. Therefore, there would be a tendency to reward oneself about as often as the model rewarded himself yielding a main effect of model's rate of self-reinforcement. This particular study appears to be amenable to a Social Comparison analysis.

The second type of research in the self-reinforcement literature concerns the adoption of specific criterion levels for self-reinforcement. In this line of research with children, investigators have utilized the miniature bowling game where a subject collects points for knocking down pins (Bandura and Kupers, 1964; Bandura and Whalen, 1966; Liebert and Ora, 1968; Mischel and Liebert, 1966). Since the end of the alley was shielded, the experimenter provided feedback to the subject which indicated how many points the subject collected on each trial. This arrangement allowed the experimenter to vary the points awarded to the subject to fit the experimental design. The model performed first and awarded himself candy (Bandura and Kupers, 1964; Bandura and Whalen, 1966) or chips to be exchanged for prizes (Liebert and Ora, 1968; Mischel and Liebert, 1966) for scores above a criterion level. Then the subject performed in the absence of the model to reduce situational pressures for conformity to the model's standard. In general, these studies found that children would adopt the standard for self-reinforcement displayed by the model. In addition, Bandura and Kupers (1964) found that adult models were more efficacious in transmitting these standards than peer models. Liebert and Ora (1968) found that modeling and direct training were equally effective in transmitting the self-reinforcement standard. They also found that there was greater violation of the standard when the prizes to be

awarded for the chips were attractive than when they were not attractive. Mischel and Liebert (1966) found that subjects would adopt a more lenient criterion for self-reinforcement if there was a discrepancy between that adopted by the model and that which the model imposed on the subject. However, when both the modeled and the imposed criteria were stringent, the subject would adopt and continue to maintain the stringent criterion. Finally, Bandura and Whalen (1966) have found that subjects will not adopt the criterion level for self-reinforcement displayed by the model if that criterion is highly divergent from the subject's own abilities as demonstrated in an earlier task. Neither the Mischel and Liebert (1966) nor the Liebert and Ora (1968) study provided a test of the Social Comparison postulates. Generally, the miniature bowling game can be seen as a perceived test of skill. The subject does not know that the experimenter is controlling the scores that he and the model receive. Moreover, the score in and of itself is meaningless. The child learns what the score he receives means in terms of reinforcement by comparing his score with that of the model. If the child's score exceeds that score for which the model has rewarded himself, the child appears to deem that score as deserving of reward. Since the subject received scores as high as those obtained by the model, there is little or no reason for the subject to infer that the model has greater competence. Therefore,



it should make little or no difference whether the model is an adult or a child. In three of these studies (Bandura and Whalen, 1966; Liebert and Ora, 1968; Mischel and Liebert, 1966), there is no test of this hypothesis since there is only one model, an adult. However, in the Bandura and Kupers (1964) study, there is a test of whether or not adult and peer models are equally effective in transmitting self-reinforcement standards.

In the Bandura and Kupers (1964) study, the children viewed either an adult model or a peer model playing the miniature bowling game. In addition, the model adopted either a high or a low criterion level for self-reinforcement. Although the subjects generally adopted the criterion level displayed, there was a more precise match in the adult model condition than in the peer model condition. Since, as argued above, there is no reason for the subject to perceive that the adult had greater ability, this finding is contrary to that predicted on the basis of Social Comparison theory. Two different lines of reasoning help to make the above contrary finding more understandable. Firstly, the adult may be perceived as more knowledgeable in terms of how many points are necessary before self-reinforcement is deserved. Were this the case, then one would expect a more precise match with the adult's standard because of the supposed greater knowledge of the adult. Secondly, it is possible that the perceived power of the adult model yielded the

greater effectiveness of the adult model in transmitting the standard. Since children model adults more than peers (Flanders, 1968), one would expect the adult model to be more efficacious. If peers were perceived to be less powerful as models, one would expect to find greater violation of peer standards. In fact, Bandura and Kupers (1964) found greater violation of the peer model's standards. The authors attempted to correct for this artifact by having the subjects perform the task alone, thus reducing situational pressures for conformity. However, if the children had learned previously that they should follow the behavior of adults more than peers due to the greater power of adults, they may adhere to the adult model's standard even in his absence. Therefore, while one can interpret the results of the Bandura and Kupers (1964) study as an argument against the Social Comparison formulation presented in this study, it is felt that other factors as mentioned above may have influenced the results. However, even though the Social Comparison predictions still appear to be tenable, it should be noted that the results of the Bandura and Kupers (1964) study may indicate that social comparison processes may not operate in all situations.

The Bandura and Whalen (1966) study was designed to be a test of the Social Comparison predictions in a self-reinforcement task. Subjects were provided with a series of three tasks which they performed. Half of the subjects were

told both at the end of each task and again at the conclusion of the series that they had performed at a level below the norm for children of their age. The other half of the subjects were informed that they had performed above the norm. This manipulation provided a success-failure experience before watching a model play the miniature bowling game. Then subjects observed either a superior model, an equally competent model, or an inferior model perform. The results indicated that the success experience subjects generally adopted the self-reinforcement standard displayed by the equally competent model. The failure experience subjects generally matched the standard of the inferior model. The superior model was ignored. Bandura and Whalen (1966) argue that this study demonstrates the influence of social comparison processes in the adoption of self-reinforcement criteria. They reason that the success-failure manipulation created an expectancy of future performance. Thus, the models which diverged from this expectancy were considered inappropriate comparison persons on the basis of ability to perform the task and rejected. While this study does provide support for the influence of social comparison processes, it does not provide a complete test of the Social Comparison predictions for two reasons. In the first place, the original success-failure manipulation was based upon performance norms for children of that age. Since norms are computed on the basis of scores for large numbers of subjects on specific

criteria, the use of norms involves an explicit comparison of the subject's score with similar other children. Therefore, there was no objective, non-social cue with which the subject could compare his performance. Social Comparison theory as developed by Festinger (1955) states that one compares himself with others when objective, non-social cues are absent. In order to constitute a complete test of the Social Comparison predictions, another group of subjects should have had the opportunity to compare their scores against some objective standard. In terms of the Bandura and Whalen (1966) study, it would have been impossible to have an objective, non-social comparison standard with the tasks utilized. The second problem, which Bandura and Whalen (1966) mention, concerns the confounding of model ability with model standards for self-reinforcement. Whenever the manipulation of model-subject similarity is accomplished by varying the ability of the model in relation to that of the subject, one runs the risk of confounding model ability with the standard he adopts. Thus, a high ability model may only perform at or above the standard he adopts or he will not be perceived as a high ability model. This also holds true for an inferior ability model. However, this also means that one is never sure if the high ability model is rejected because he is much too able as predicted by Social Comparison theory or because the standard adopted by the high ability model is too stringent. In order to



provide an adequate test of the predictions based upon Social Comparison theory, one must separate out this confounding. It should be pointed out that Bandura and Whalen (1966) noted this problem, but they felt that this confounding was not important to their study. However, it is important to a study designed to interpret self-reinforcement effects in terms of Social Comparison predictions.

### Proposed research and hypotheses

This particular study attempted to provide an adequate test of Social Comparison predictions for self-reinforcement effects. On the basis of Social Comparison theory it was argued that the similar model's self-reinforcement behavior would be adopted more than the behavior of the dissimilar model when there were social cues available. However, there should be no differences in amount of modeling between the similar and dissimilar model conditions when the available cues are objective and non-social. Subjects were asked, after watching a model, to perform a prediction task under two different sets of task instructions. In one condition, subjects were told that the order of illumination of two lights, which the subject was to predict, followed a random order. In this condition, it was argued that subjects could compare their performance with chance and select self-reinforcement appropriately. When the illumination of the lights was random, the subject's predictions were either

right or wrong. Therefore, subjects might select either all the self-reinforcement available or none at all depending upon whether the subject was correct or incorrect. In the second task instruction condition, subjects were told that the order of illumination of the lights followed a definite pattern. It was argued that subjects would attempt to discover the pattern in the illumination of the lights. While looking for this pattern, subjects would have difficulty in determining how much self-reinforcement was appropriate. Therefore, these subjects would tend to match the self-reinforcement behavior of the model. The dependent measure of amount of self-reinforcement was intended to provide some measure of the subject's evaluation of his performance.

In addition, the confounding of model ability with model standard was separated. Instead of manipulating the ability discrepancy between the model and the subject, similarity of attitude concerning the importance of doing well on the task was manipulated. Berger (1971) has demonstrated that varying the similarity of attitude between the model and the subject is sufficient to involve social comparison processes. Subjects in the model similar condition were given feedback that the model agreed with their position concerning the importance of doing well on the task while subjects in the dissimilar model condition received feedback to the effect that the model disagreed with their position. Interpretation of Berger's (1971) data would argue that

similarity in either ability or opinion qualifies the comparison other as an appropriate comparison person.

## METHOD

### Subjects

Subjects ( $N = 64$ ) were male and female students at the University of Massachusetts, Amherst, who received extra course credit for participation in psychological experiments. Four of the subjects were dropped from the study before analysis of the data either due to failure to understand the experimental instructions or due to mechanical failure of the apparatus. One subject each was removed from the dissimilar model, patterned order and the similar model, random order conditions and two subjects were removed from the similar model, patterned order condition. The remaining sixty subjects were randomly assigned to experimental conditions, yielding fifteen subjects per cell. All subjects were run individually.

### Procedure

Design. The study was a  $2 \times 2 \times 10$  experimental design with task instructions (patterned order versus random order) and model similarity (similar versus dissimilar) as the between-subjects variables and correct trials as the within-subjects variable.

Apparatus. The subject was shown a panel consisting of four lights arranged in two rows of two lights, with two buttons below the bottom row of lights. The top blue row was labeled feedback and the bottom red row was labeled prediction. The subject was required to predict which of the two lights in the feedback row would be illuminated. His prediction was made by pushing one of the two buttons which illuminated the corresponding light in the prediction row. After a delay of five seconds one of the lights in the feedback row was illuminated. If the light in the same column as his prediction was illuminated, he was correct. However, if the other light was illuminated, he was incorrect. While the panel was ostensibly connected to a relay rack containing timers, relays, and a Talley tape reader, the illumination of the feedback lights was under the experimenter's control.

Experimental sequence. After escorting the subject to the experimental room, the experimenter introduced himself and read the following instructions:

The study you have selected is concerned with personal motivation. Social scientists have long been concerned with how an individual maintains interest and attention while engaged in a repetitive task. For example, assembly line workers and air traffic controllers are required to repeat the same task over and over. If an air traffic controller loses interest in his job, an unidentified airplane may enter his sector and a fatal crash might result. It is known that some air traffic controllers praise or criticize themselves to maintain interest in their job. For example, an air traffic controller may praise himself whenever he spots an unidentified plane entering his sector. Or he may criticize himself whenever he fails to



identify a plane as it enters his sector. Some air traffic controllers use points instead of praise and then compare one day's point total with another day's point total. This particular study is an attempt to see if techniques such as giving points really are effective in increasing attention to a task.

What I would like you to do is predict which of the two lights in the top row of the panel before you will be illuminated. If you will look at the panel, you will see that there are two rows of two lights. The top blue row is labeled FEEDBACK and the bottom red row is labeled PREDICTION. Below the bottom row of lights there are two buttons. Whenever one of the buttons is pushed, the light in the PREDICTION row above that button will be illuminated indicating your choice. Please push the button quickly and then release. Never hold the button down. About five seconds after pushing either button, one of the lights in the FEEDBACK row will be illuminated. If the light in the FEEDBACK row above the PREDICTION light is illuminated, then you were correct. However, if the other light is illuminated, then you were incorrect. Notice that there is an amber light at the top center of the panel. This is a warning light. I control this light and it is a signal for the trial to begin.

At the conclusion of each trial, please tell me how many points you feel you deserve for your performance on that trial. You may give yourself as few or as many points as you feel you deserve. However, there is one restriction. You may not give yourself more than fifteen points for any trial. Therefore, you may give yourself from 0 to 15 points for any trial.

The electronic equipment behind you controls the order in which the lights in the FEEDBACK row will be illuminated.

At this point the experimenter showed the subject a piece of punched paper tape. Then, the patterned versus random order task instruction manipulation was introduced. In the patterned order condition, the experimenter went on to state:

This tape will program the apparatus so that the illumination of the FEEDBACK lights follows a

definite pattern. While it is a complex pattern, you should be able to notice it.

In the random order condition, the experimenter stated:

This tape will program the apparatus so that the illumination of the FEEDBACK lights is completely random. Therefore, if you made enough predictions you would be correct 50% of the time. However, in a brief series of trials you may be either above or below this chance level depending on luck.

Then the instructions for all subjects continued as follows:

Remember, first you will see the amber light. When the light goes off, select either button, push quickly and release. The illumination of the light in the FEEDBACK row will indicate whether or not your prediction was correct. Finally, tell me how many points from 0 to 15 you deserve for that trial. Remember, only give the number of points you feel you deserve for your performance on that trial.

So that you will get an idea of how the apparatus works, I am going to show you an excerpt of a video tape of an earlier subject who went through the same procedure. However, before we start I would like you to fill out this scale. We have found that a subject's performance on this type of task is influenced by how important he feels the task is. Please indicate whether or not you feel that doing well on this task is important to you.

The experimenter then gave the subject a questionnaire consisting of a single statement. The question asked the subject if it was important to him that he do well on the experimental task. The subject indicated his response by checking either "yes" or "no." While the subject was completing the scale, the experimenter readied the equipment. When the subject had finished, the experimenter picked up the scale and casually remarked in the similar model condition:

That's interesting. The subject you are going to see marked the same answer.

In the dissimilar model condition, the experimenter said:

That's interesting. The subject you are going to see marked the opposite answer.

At this point the subject viewed a female model perform a series of 10 trials on the experimental task. When she was correct, the model selected seven points for her self-reinforcement. However, when she was incorrect, she did not give herself any points. After the subject had viewed the model, the sequence constituting a trial was reviewed, the subject asked if he had any questions, and the experiment began. The subject completed a series of twenty trials, on half of which he was given feedback that he was correct. With I signifying incorrect and C signifying correct, the following sequence constituted the subject's feedback: I, I, C, I, I, C, I, I, C, I, C, I, C, C, C, I, C, C, C, I. After completing this series of twenty trials, the subject was asked to fill out a final questionnaire. When the questionnaire was returned, subjects were interviewed to see if they had discerned the purposes of the study, fully debriefed as to the purposes and manipulations of the study, thanked for their participation, given a hand-out describing the study, and dismissed.

Questionnaire. A questionnaire (see Appendix A) was designed to check on the operation of the independent manipulations and tap possible alternative explanations for the

results of the study. Part of the questionnaire consisted of six seven-point rating scales. The first scale asked whether or not the self-administration of points affected the subject's attention in the experimental task. The end points of the scale were labeled 1 (increased attention) and 7 (decreased attention). The purpose of this scale was to lend credence to the cover story that the study was, in fact, concerned with whether or not the administration of points can affect attention. The second scale asked subjects how they would rate their performance on the basis of chance. The end points of this scale were labeled 1 (below chance) and 7 (above chance). It was hypothesized that subjects in the random order conditions would rate their ability at chance. However, subjects in the patterned order condition, if the manipulation was successful, should rate their ability as better than chance. This argument is based on the premise that subjects who felt that they had discerned a pattern in the sequence of illumination of the feedback lights should feel that they were performing at better than a chance level. The third question asked subjects whether the other subject (model) differed in ability from the subject. The end points of this scale were labeled 1 (very different) and 7 (not very different). This question was included to check on a possible equity theory (Adams, 1967) interpretation of the data. Masters (1968) has noted that self-reinforcement can be used to maintain a proportional



ratio between inputs into a task and payoffs one receives if the subject perceives his performance as different from another person's performance. The fourth question provided a manipulation check on the operation of the model similarity manipulation. It asked the subject how similar in attitude was the subject they watched in her rating of the importance of doing well on the task. The end points of this scale were labeled 1 (quite similar) and 7 (quite dissimilar). The final rating scale asked how important doing well had become on the task after having completed the task. The end points of this scale were labeled 1 (very important) and 7 (not very important). This scale was included to determine whether or not the subjects' subjective evaluations of the importance of the task changed as a function of their experience with the task. Such changes could alter the operation of the model similarity manipulation. The questionnaire also asked the subjects to state how many points the model gave herself when correct in order to determine if the subjects had understood the model's behavior. Subjects were also asked if they saw a pattern to the sequence of illumination of the feedback lights to assess the operation of the task instructions manipulation. Subjects responded to this question by checking either a "yes" or a "no." Open-ended questions asked subjects if other techniques might be more effective in increasing attention in a repetitive task, how they determined the amount of self-

reinforcement they deserved, and to indicate what they thought the experiment was all about.

Statistical analyses. Inspection of the data indicated that, contrary to expectations, subjects frequently rewarded themselves below as well as above the model's standard. Obviously, self-reinforcement below the model's standard is as much a rejection of the model's standard as self-reinforcement above her standard. However, simply averaging the data would lead to a situation where the mean would approximate the standard displayed by the model. Therefore, it was necessary to utilize some type of transformation which would more accurately reflect acceptance or rejection of the model's self-reinforcement standard. With this in mind, it was decided to use an absolute deviation from the model's standard as the major dependent measure. These data were then analyzed by means of an analysis of variance. The scaled questionnaire data were also analyzed by means of the analysis of variance procedure.

## RESULTS

### Self-reinforcement data

Before reporting the results of the analysis of variance for the self-reinforcement data it is necessary to add one caution. Plotting the frequency of the scores, it was found that there was a bimodal distribution of the data. This bimodality appeared to be caused by subjects in each condition opting to use the maximum number of points possible

whenever they were correct. Unfortunately there was no transformation available which would correct for this problem as well as no non-parametric statistics appropriate to the data. Therefore, it was decided to utilize an analysis of variance since this procedure is particularly robust (Myers, 1967). However, interpretation of the results of this analysis of variance should be treated with caution.

Table 1 reports the results of the analysis of variance computed for the deviation score data. Inspection of Table 1 indicated that the only significant effect was that

-----  
 Insert Table 1 about here  
 -----

of trials ( $F = 6.12$ ,  $9/504$  df,  $p < .001$ ). The main effects of model similarity and task instructions (patterned versus random order) as well as all interactions failed to reach statistical significance.

Table 2 presents the cell means for all conditions. The results of a  $t$ -test for correlated scores indicated a significant decrease ( $t = 3.57$ ,  $df = 59$ ,  $p < .001$ , two-tailed) in modeling from the first correct trial ( $\bar{X} = 3.82$ ) to the last correct trial ( $\bar{X} = 4.88$ ).

-----  
 Insert Table 2 about here  
 -----

Inspection of Table 2 indicates that the ordering of the overall means for the between-subjects cells tended to support the experimental hypotheses for the patterned order groups. That is to say, within the patterned order task instruction condition, the similar model group showed more

TABLE 1  
SUMMARY OF ANALYSIS OF VARIANCE FOR ABSOLUTE  
DEVIATION FROM THE MODEL'S STANDARD

SV	df	MS	F
Between <u>Ss</u>			
A (model similarity)	1	22.43	<1
B (task instructions)	1	39.53	<1
AB	1	158.11	2.38
S/AB	56	66.39	
Within <u>Ss</u>			
T (trials)	9	10.06	6.12*
TA	9	1.71	1.04
TB	9	.57	<1
TAB	9	1.16	<1
ST/AB	504	1.64	

\*p < .001



TABLE 2  
CELL MEANS BY CONDITIONS FOR SELF-REINFORCEMENT DATA<sup>1</sup>

Instructions	Model Similarity	Trial Numbers for the Correct Trials										Mean
		3	6	9	11	13	14	15	17	18	19	
Patterned Order	Similar	3.13	3.13	3.26	3.33	3.06	3.80	4.53	3.80	4.27	4.93	3.73
	Dissimilar	5.13	5.00	5.20	5.20	4.93	5.07	5.33	4.93	5.47	5.33	5.14
Random Order	Similar	3.87	3.80	4.00	3.53	4.00	4.53	4.93	4.13	4.73	4.87	4.24
	Dissimilar	3.13	3.33	3.53	3.33	3.47	3.47	4.33	3.47	3.53	4.40	3.60
Mean		3.82	3.82	4.00	3.85	3.82	4.22	4.78	4.08	4.50	4.88	

<sup>1</sup>The lower the number, the more the modeling

modeling ( $\bar{X} = 3.73$ ) than the dissimilar model group ( $\bar{X} = 5.14$ ). However, the ordering of the means within the random order task instruction condition was exactly the opposite. In the random order condition, the dissimilar model group showed more modeling ( $\bar{X} = 3.60$ ) than the similar model group ( $\bar{X} = 4.24$ ). Generally, the ordering of the between-subjects means are in line with the experimental hypotheses for the patterned order condition but not for the random order condition. However, it should be noted that none of the differences among the means were statistically significant.

#### Questionnaire data

The questionnaire data (see Appendix A) were analyzed by means of separate analyses of variance for each question. The data from the questions concerning the effect of self-reinforcement on attention, the subject's ability in terms of chance, the difference in ability between the model and the subject, and the importance of doing well on the experimental task all failed to yield statistically significant results. Table 3 presents a summary of the analyses of variance for the questionnaire data. It is important to note that many of the  $F$ s for the non-significant data were less

-----  
 Insert Table 3 about here  
 -----

than 1.00, indicating that the variables these questions were designed to tap probably were not affected by the experimental manipulations.

TABLE 3  
ANALYSES OF VARIANCE FOR QUESTIONNAIRE DATA

SV	df	Effect of self-reinforcement on attention		S's ability in terms of chance		Difference between $\bar{S}$ and $\bar{M}$ in ability		Helpfulness of $\bar{M}$ when making predictions		Model Similarity		Importance of doing well on the task	
		MS	F	MS	F	MS	F	MS	F	MS	F	MS	F
A (model similarity)	1	1.07	<1	.82	<1	6.67	1.91	1.35	<1	129.07	2782**	.15	<1
B (task instructions)	1	.07	<1	1.35	1.01	6.67	1.91	.42	<1	5.40	1.16	3.75	1.33
AB	1	.07	<1	.82	<1	.27	<1	16.02	4.19*	3.27	<1	.82	<1
S/AB	56	1.82		1.34		3.49		3.82		4.64	2.81		

\*p < .05

\*\*p < .001

The data from the question asking whether or not the model was helpful to the subject in making the predictions yielded a significant interaction between model similarity and task instructions ( $F = 4.19$ ,  $1/56$  df,  $p < .05$ ). Table 4 presents the cell means for this data.

TABLE 4  
CELL MEANS FOR HELPFULNESS OF MODEL DATA<sup>1</sup>

	Patterned Order Instructions	Random Order Instructions
Similar Model	3.80	2.60
Dissimilar Model	3.07	3.93

<sup>1</sup>The larger the number, the more helpful the model was to the subject when the subject made his predictions.

Inspection of this table indicates that in the patterned order condition the model was seen as more helpful when the model was described as similar as opposed to dissimilar ( $\bar{X} = 3.80$ , model similar, patterned order;  $\bar{X} = 3.07$ , model dissimilar, patterned order). This ordering is reversed for the random order condition. In this condition the model was seen as more helpful to the subject in making his predictions when the model was dissimilar ( $\bar{X} = 3.93$ ) than when the model was seen as similar ( $\bar{X} = 2.60$ ). This interaction was subsequently analyzed by means of Duncan's

Multiple-Range Test (Duncan, 1955). The results of this analysis indicated that within the similar model groups the patterned order task instruction condition differed from the random order task instruction condition ( $p < .05$ ). In addition, it was also found that the similar model group differed from the dissimilar model group in the random order task instruction condition ( $p < .05$ ).

Questions designed to assess the operation of the independent variable manipulations indicated that the model similarity manipulation was successful while the patterned versus random order manipulation was unsuccessful. There was a strong main effect of model similarity ( $F = 27.82$ ,  $1/56$  df,  $p < .001$ ) with the similar model perceived as similar in attitude concerning the importance of doing well on the task ( $\bar{X} = 2.80$ ) and the dissimilar model perceived as dissimilar ( $\bar{X} = 5.73$ ). In order to test whether or not the task instructions (patterned versus random order) manipulation was successful, subjects were asked if they had noticed a pattern to the illumination of the feedback lights. Since the response to this question involved checking either "yes" or "no," the data were inappropriate for an analysis of variance. Therefore, a  $\chi^2$  analysis was conducted using the frequency of "no" responses as the dependent variable. The results indicated that there were no differences among the four experimental groups ( $\chi^2 = .4518$ ,  $df = 1$ ). Thus, on the basis of this question it was concluded that the task instructions



manipulation failed to operate as anticipated.

Finally, responses to the question designed to ascertain whether or not the subjects understood the amount of self-reinforcement selected by the model indicated that subjects were aware of how many points the model selected.

## DISCUSSION

While the results of the analysis of variance conducted on the self-reinforcement data failed to yield any significant effects of either model similarity or task instructions (patterned versus random order), the ordering of the cell means indicated some support for the experimental hypotheses. The hypotheses were supported in the patterned order condition with the similar model group modeling the self-reinforcement behavior of the model more than the dissimilar model group. In the random order condition, it was predicted that there would be no differences between the model similarity groups. However, the trends appeared to indicate that the dissimilar model group showed more modeling than the similar model group. Nevertheless, the differences among these means were non-significant. Overall, the only significant effect in the self-reinforcement data was the main effect of trials. The questionnaire data indicated that the model similarity manipulation operated as expected while the task instruction (patterned versus random order) manipulation did not. Other questionnaire data indicated that the

independent variables did not differentially affect either the subjects' perceptions of the influence of self-reinforcement on attention, the subjects' perceptions of their abilities in terms of chance, the subjects' perceptions of any differences in ability between themselves and the model, or the subjects' perceptions of the importance of doing well on the task. However, data concerning whether or not the model was perceived as helpful to the subject in making his predictions yielded a significant interaction between model similarity and task instructions. This interaction indicated that the similar model was perceived as more helpful than the dissimilar model in the patterned order conditions while the dissimilar model was perceived as more helpful than the similar model in the random order conditions.

Due to the lack of significant effects, it is impossible to either support or disconfirm the experimental hypotheses which prompted this study. Obviously, it can be argued that the failure of the instructions manipulation negated any hope for an adequate test of the Social Comparison formulations, and thus, inconclusive results should be expected. However, three lines of reasoning argue against this position. Firstly, the patterning of the means was generally in line with the hypotheses except for the random order, model dissimilar condition. Due to the high variability of the data, it can be argued that the lack of

significant effects was due to too few subjects rather than the instructions manipulation. Possibly future research using a prediction task should use more subjects than were used in this study.

Secondly, when subjects were regrouped on the basis of the questionnaire data to yield groups which perceived the independent variable manipulations appropriately, the patterning of the means was parallel to that of the total data. In other words the similar model group modeled more than the dissimilar model group when the task instructions specified a patterned order to the illumination of the lights. In the random order condition, the dissimilar model group showed more modeling than the similar model group. Since the patterning of the cell means for the regrouped data paralleled that of the total data, it is possible that the total data reflected the actual behavior of the subjects rather than a failure of the task instructions manipulation. Unfortunately, when the data were regrouped, the resulting number of subjects was too small for statistical analysis.

Finally, it is possible that the failure of the task instructions manipulation check to yield significant results was due to the wording of the question itself. The pilot data indicated that a single statement asking subjects whether or not they had seen a sequence to the illumination of the lights was sufficient to differentiate the patterned and random order task instructions conditions. In the study

itself, no such difference was found. However, subjects reported in the post-experimental interview that they had searched for a pattern but were unable to find one. From a Social Comparison position, belief in and search for a pattern during the series of trials should be sufficient for the arousal of social comparison processes. In sum, it is argued that the failure of the task instructions manipulation check may not have reflected a failure of the task instructions manipulation. The reasons for this argument include (1) the patterning of the overall between-subjects means both before and (2) after regrouping the data, and (3) the possibly inappropriate wording of the task instructions manipulation check.

The trials main effect ( $F = 6.12$ ,  $1/504$  df,  $p < .001$ ) holds some interest. Further data analysis of the trials data indicated a significant ( $t = 3.57$ ,  $df = 59$ ,  $p < .001$ , two-tailed) change in amount of self-reinforcement administered with more modeling on the first trial ( $\bar{X} = 3.82$ ) than the last trial ( $\bar{X} = 4.88$ ). There are two plausible explanations for this trials effect. Firstly, it may be argued that subjects confronted with this particular prediction task initially opted to follow the model and then after some experience decided to develop their own criteria for self-reinforcement. The second argument simply states that subjects became bored over the series of trials, and thus, opted to ignore the self-reinforcement behavior of the model. Since



it is impossible to distinguish between these possible explanations on the basis of the data, the trials main effect remains largely unexplained.

Frequently, post-experimental questionnaire data can provide important clues as to why a particular study either succeeded or failed. In addition, questionnaire data often provides evidence concerning the plausibility of alternative explanations of the data. Unfortunately, such is generally not the case in this particular study. Nevertheless, the data from the question asking whether or not the model was helpful to the subject when the subject made his predictions held some interest. As noted earlier, the means for this scale indicated that the similar model was more helpful than the dissimilar model for subjects receiving the patterned order task instructions. However, the opposite relationship held for the random order task instructions condition. This patterning of means was similar to that found for the between-subjects variables of model similarity and task instructions in the self-reinforcement data. It is possible that the subjects who found the model helpful when subsequently making their own predictions also found the model helpful when determining how much self-reinforcement was appropriate. In order to test such a hypothesis, a correlation between perceived helpfulness of the model and amount of self-reinforcement selected was computed. Due to the nature of the rating scale with higher numbers indicating greater

helpfulness and lower self-reinforcement scores indicating greater modeling, a negative correlation was necessary to support the hypothesis. Unfortunately, the obtained correlation was  $-.18$  ( $df = 58$ ,  $p > .05$ ) indicating no relationship between perceived helpfulness of the model and amount of self-reinforcement administered.

Since the questionnaire data failed to provide any significant clues as to why the independent variables failed to have a significant effect on the amount of self-reinforcement subjects gave to themselves, it is necessary to seek possible explanations from two other sources. Firstly, one must re-examine Social Comparison theory in order to determine whether or not the experimental operations of the study adequately fulfilled the conceptual formulations of the theory. Secondly, one must re-examine the nature of the task utilized to see if there might be some aspect of the task which might have contributed to the obtained results.

Festinger (1968) has discussed his conceptualization of the social versus non-social standard which an individual selects in order to determine the appropriateness of his opinions and/or abilities. According to Festinger, this dimension may be viewed as having some type of physical reality at one end and a social reality at the other. An attitude concerning a physical reality may be empirically tested. As an example, Festinger noted that if one felt that the ice covering a pond was too thin, one could test the thickness

of the ice by a number of techniques. In other words, one may empirically determine the correctness of his attitude by actually testing it in the physical world. At the other end of the dimension is social reality which cannot be empirically tested. For example, one may believe that a given examination was too difficult after failing it, but there is no real physical test of whether or not the examination was too difficult. In this situation, Festinger argues, one seeks out others who share this belief in order to gain some additional support for the belief. In the present study, it was argued that the instructions stating that the order of illumination of the feedback lights was random constituted a physical reality by which the subject could determine the adequacy of his performance. This physical reality was chance performance. Since the subject knew that each trial was a guess, it was argued that the subject would have no reason to compare himself with the model in order to determine how many points a correct response was worth. On the other hand, subjects in the patterned order conditions were told that the illumination of the feedback lights followed a definite pattern. It was argued that these instructions placed the subject in a situation where he had to perform a series of trials before he could decide whether or not he had discovered the pattern. However, while searching for the pattern, the subject would have no standard available to determine how many points a single trial was worth. Therefore,

it was argued that a subject in this situation would rely on the behavior of the model in order to determine how many points each individual trial was worth. The trends in the self-reinforcement data noted earlier appear to support the idea that the task instructions reflect Festinger's social-physical reality dimension. On the other hand, since there were no significant differences among the overall between-subjects group means, it can be argued that the trends merely reflect random error. Such a possibility and its implications both for this particular experimental design and future designs testing Social Comparison predictions in the modeling of amount of self-reinforcement must be considered. Since the between-subjects means were not statistically different from one another, it is possible that the task instructions failed to provide a physical-social reality. Specifically, it could be argued that the patterned order instructions failed to provide a social reality as Festinger (1955) conceptualized it. In other words, subjects in the patterned order condition might have felt that the self-reinforcing behavior of any model (similar or dissimilar) was irrelevant since they were testing whether or not their predictions fit into a pattern purportedly determined by some electrical apparatus. Thus, it may be that the patterned order instructions actually reflected a degree of physical reality rather than a social reality. If this were the case there would be no differences between the task instruction (patterned versus



random order) conditions since both would reflect a physical reality. In addition, the presence or absence of a model would be irrelevant since a comparison other is not considered when a physical reality is present. Overall, there should then be no differences among the between-subjects means. Obviously, if the experimental manipulations do not accurately reflect the concepts of Social Comparison theory, these manipulations can operate in the expected manner without providing a test of the theoretical formulation. Naturally, such a problem was not foreseen in the design of the experiment. If the above explanation is accepted, then this study did not provide a test of the Social Comparison theory predictions. However, before such an explanation is accepted, one caution should be mentioned. If the instructions manipulation had merely reflected a physical reality instead of both a physical and a social reality, one would not expect the model to have any effect on the selection of self-reinforcement. However, data from the pilot research indicated that subjects in all model conditions exhibited significantly more modeling of the model's amount of self-reinforcement than did subjects in a no model control condition. Regardless, any difference between model conditions and no model controls may be ascribed to some type of general modeling influence. Therefore, future research should investigate whether or not task instructions concerning the illumination of a set of lights really reflects Festinger's (1968) physical-social reality dimension.

As mentioned above, the second source of clues as to why the independent variables failed to influence the amount of self-reinforcement selected concerns the task itself. The above discussion alludes to problems with the task. However, its emphasis is on whether or not the experimental instructions accurately reflect the conceptualizations of the social-physical reality dimension. Before examining the task, however, it is important to point out the difficulties inherent in developing a single task which can be varied along a physical-social reality dimension. After much reflection, it was determined that the only objective, non-social standard feasible was in terms of probability. Chance is a mathematical concept which is empirically determined and has no social referent. In addition, a probability task such as the one used can be varied so that the standard can be changed from chance instructionally to one which is no longer chance by having a subject search for a pattern in the illumination of lights. When looking for a pattern, there is also an empirically determined standard as discussed above; whether or not one finds the pattern. In addition, one can easily introduce a social referent (a model) whose similarity to the subject can be varied along an attitudinal dimension. Thus, it was argued that a probability task could be modified in such a way as to reflect the extremes of the social versus objective, non-social standard dimension as stipulated by Festinger (1955) as well as varying the similarity of the model.

Unfortunately, the use of a probability type of task has a serious disadvantage. In the minds of many subjects a probability task such as the one used utilizing predictions of an event is perceived as simply guessing. When one guesses, one is right or wrong. If one is right, he deserves something and if one is wrong he doesn't. Numerous strategies enter into the situation when one is determining how much he deserves when correct or incorrect. One simple strategy was to award oneself all the points when correct and none when incorrect since one is just guessing. Another more elaborate strategy involved giving oneself half the number of points possible on any correct trial since one had a 50-50 chance of being correct on any given trial. Regardless of the specific strategy used to determine the amount of self-reinforcement, the use of strategies implies a failure to transmit to the subject the fact that he can rely on the model for cues as to appropriate self-reinforcement. It is possible that the use of a prediction task implied some type of guessing even though the subject received instructions to the contrary. Possibly only a prediction task runs the risk of being reduced, in the subject's mind, to merely guessing. If that is the case, then future research could possibly investigate other types of tasks to see if they fulfill the requirements of Festinger's (1968) physical-social reality dimension and still be amenable to the introduction of a model.

When research fails to provide clear cut answers to experimental hypotheses, it is incumbent upon the researcher to point out directions for future research. The ideas presented in this paper point to two directions for future research. The first direction would be to conduct research using various tasks which reflect Festinger's (1955) physical-social reality. Since it is difficult to develop a single task which reflects this dimension, it is possible to utilize a series of tasks which reflect the end points of the dimension. If such research indicated that modeling of self-reinforcement does not vary as a function of task, it would be possible to argue that any obtained differences among the cell means actually reflect the operation of the independent variables separate from any specific task. The second approach to the problem of developing tasks which reflect the physical-social reality dimension would be to utilize a normative standard as the physical reality. Bandura and Whalen (1966) utilized such an approach. However, in this paper it was argued that the normative standard implies some type of social comparison since a social standard is a statistical average of all individuals along a particular dimension. However, it is possible that the concept of norm has evolved in such a way that many consider a norm to be a standard. If it can be empirically demonstrated that a normative standard is considered to be some type of physical reality, then its utilization in research designed to test



Social Comparison predictions for the modeling of self-reinforcement would be appropriate. Nevertheless, without such an empirical demonstration, which is the case with the Bandura and Whalen (1966) study, the use of a normative standard as a physical reality is inappropriate.

In sum, although the trends of the data partially supported the hypotheses which prompted this study, the lack of significant results indicate that much more research is necessary before one can state whether or not social comparison processes are important in the modeling of rate, standard, or amount of self-reinforcement. Suggestions for future research on this problem center around a clearer specification of the social-physical reality dimension and the types of tasks used to operationalize this dimension. As Bandura (1971) has noted, the role of self-reinforcement in the possible maintenance of deviant responses as well as the possible role of self-reinforcement in resistance to extinction make self-reinforcement an important topic for future research.

## REFERENCES

- Adams, J. S. Inequity in social exchange. In L. Berkowitz (Ed.), Advances in experimental social psychology. New York: Academic Press, 1965. Pp. 267-299.
- Bandura, A. Vicarious- and self-reinforcement process. In R. Glaser (Ed.), The nature of reinforcement. New York: Academic Press, 1971. Pp. 228-278.
- Bandura, A. and Kupers, C. J. Transmission of patterns of self-reinforcement through modeling. Journal of abnormal and social psychology, 1964, 69, 1-9.
- Bandura, A. and Perloff, B. Relative efficacy of self-monitored and externally imposed reinforcement systems. Journal of personality and social psychology, 1967, 7, 111-116.
- Bandura, A. and Whalen, C. K. The influence of antecedent reinforcement and divergent modeling cues on patterns of self-reward. Journal of personality and social psychology, 1966, 3, 373-382.
- Berger, S. M. Observer perseverance as related to a model's success: A social comparison analysis. Journal of personality and social psychology, 1971, 19, 341-350.
- Duncan, D. B. Multiple range and multiple F tests. Biometrics, 1955, 11, 1-42.
- Festinger, L. A theory of social comparison processes. In A. P. Hare, E. F. Borgatta, and R. F. Bales (Eds.), Small groups. New York: Knopf, 1955. Pp. 163-187.
- Festinger, L. Informal social communication. In D. Cartwright and A. Zander (Eds.), Group dynamics. New York: Harper and Row, 1968. Pp. 182-191.
- Flanders, J. P. A review of research on imitative behavior. Psychological bulletin, 1968, 39, 537-544.
- Gewirtz, J. L. The roles of overt responding and extrinsic reinforcement in "self-" and "vicarious-reinforcement" phenomena and in "observational learning" and imitation: Discussion of Professor Bandura's paper. In R. Glaser (Ed.), The nature of reinforcement. New York: Academic Press, 1971. Pp. 279-309.

- Hilgard, E. R. and Bower, G. H. Theories of learning. New York: Appleton-Century-Crofts, 1966.
- Kanfer, F. H. and Marston, A. R. Determinants of self-reinforcement in human learning. Journal of experimental psychology, 1963, 66, 245-254.
- Liebert, R. M. and Ora, J. P., Jr. Children's adoption of self reward patterns: Incentive level and method of transmission. Child development, 1968, 39, 537-544.
- Marston, A. R. Imitation, self-reinforcement, and reinforcement of another person. Journal of personality and social psychology, 1965, 2, 255-261.
- Marston, A. R. and Kanfer, F. H. Human reinforcement: Experimenter and subject controlled. Journal of experimental psychology, 1963, 66, 91-94.
- Masters, J. C. Effects of social comparison upon subsequent self-reinforcement behavior in children. Journal of personality and social psychology, 1968, 10, 391-401.
- Mischel, W. and Liebert, R. M. Effects of discrepancies between observed and imposed reward criteria on their acquisition and transmission. Journal of personality and social psychology, 1966, 3, 45-53.
- Myers, J. L. Fundamentals of experimental design. Boston: Allyn and Bacon, 1967.

## APPENDIX A

### Post-experimental Questionnaire



## QUESTIONNAIRE

How did awarding points affect your attention in the task?

:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:  
 1 2 3 4 5 6 7  
 increased no decreased  
 attention change attention

Do you think that other techniques might be more effective in maintaining your attention in a repetitive task? (If "yes," please specify.)

In terms of chance performance, how would you rate your ability in this particular task?

:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:  
 1 2 3 4 5 6 7  
 below chance above  
 chance chance chance

In terms of ability in performing the task, how different was the subject you watched from you?

:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:  
 1 2 3 4 5 6 7  
 very average not very  
 different average different

How much did watching the other subject help you in making your predictions?

:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:  
 1 2 3 4 5 6 7  
 not at neutral quite a  
 all neutral lot

According to the information I gave you, how similar was the other subject in her rating of the importance of doing well on this task?

:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:\_\_\_\_:  
 1 2 3 4 5 6 7  
 quite neutral quite  
 similar neutral dissimilar

Having completed the task, how important did doing well on this task become?

: 1 : 2 : 3 : 4 : 5 : 6 : 7 :  
 very                      average                      not very  
 important    important

How many points did the other subject give herself when she was correct?

\_\_\_\_\_ points

Did you see a pattern in the sequence of illumination of the lights in the FEEDBACK (top) row?

\_\_\_\_\_ yes

\_\_\_\_\_ no

How did you determine how many points were appropriate when you gave yourself points?

Please indicate what you think this experiment was all about.



