

Examining the Impact of Impression Management Context and
Self-Monitoring on the Leniency and Accuracy of Self-
Appraisals

by

Joshua Holbrook Williams

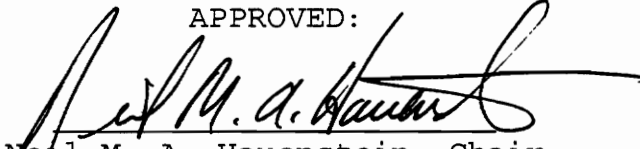
Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

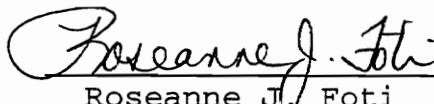
MASTER OF SCIENCE

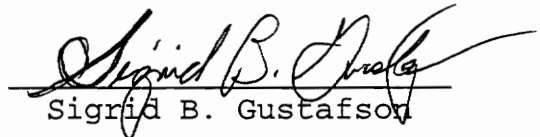
in

PSYCHOLOGY

APPROVED:


Neil M. A. Hauenstein, Chair


Roseanne J. Foti


Sigrid B. Gustafson

December 14, 1996

Blacksburg, Virginia

Keywords: Self-Appraisals, Impression Management Context,
Self-Monitoring, Self-Deception

C.7

LD
5655
V855
1996
W5548
C.2

EXAMINING THE IMPACT OF IMPRESSION MANAGEMENT CONTEXT AND
SELF-MONITORING ON THE LENIENCY AND ACCURACY OF SELF-
APPRAISALS

by

Joshua Holbrook Williams

Committee Chairperson: Neil M. A. Hauenstein

Psychology

(ABSTRACT)

Self-appraisals of performance are traditionally lenient and inaccurate, hampering their practical utility in applied settings. The purpose of the current study is to examine the underlying processes, namely self-deception and impression management, which contribute to this leniency and inaccuracy. Because self-ratings are inflated regardless of environmental affordances, self-deception is said to occur. However, when environmental contingencies that reward positive self-evaluations exist, leniency and inaccuracy increases. This suggests that impression management processes also contribute to inflated and inaccurate self-appraisals. The environmental affordances associated with self-ratings are often couched in terms of reward and non-reward purposes of appraisal (POA). The occurrence of leniency and inaccuracy in reward purposes of appraisal are potentially moderated by personality variables such as self-monitoring (SM). Consequently, POA and SM were examined in

the current study. Participants completed a model building task in both non-reward and reward POAs, with self-appraisals following each task. They also completed surveys which assessed their levels of self-monitoring, self-deception, and impression management. It was predicted that self-rated performance would be lenient across conditions, reflecting self-deception. It was further predicted that participants would be more lenient and less accurate in the reward POA than in the non-reward POA, reflecting impression management processes. This would suggest an additive effect in which impression management leads to increased inflation beyond the level of inflation attributed to self-deception. Finally, it was predicted that self-rating leniency in the reward POA would be moderated by self-monitoring, such that only high self-monitors would be significantly more lenient in the reward POA in terms of their self-rated performance. Repeated measure ANOVAs using four accuracy and four leniency measures yielded limited support for the hypotheses. Implications for future research are discussed.

ACKNOWLEDGMENTS

This document is dedicated to my Mom and Dad. They have given me unconditional love and support my whole life. I love you guys very much and am lucky to have you as my parents.

I thank Dr. Neil Hauenstein for chairing my thesis and seeing me through this project. I would also like to thank committee members Dr. Roseanne Foti for her support and professionalism and Dr. Sigrid Gustafson for her support and offbeat sense of humor.

I also thank CC for her love and belief in me. The skies cleared when she came into my life. In addition, I thank my great friends who I've been graced to have here: Bobo for her kindness, Niki and Jason for their sense of perspective, Theresa for her smile, and Dan for just being Dan. Dan went beyond the call of duty in being a best friend with both technical and emotional support. Also, Jeff, Thomas, Heath, Steve, Bobby, Jason M., Super G, and Chuck have also been great friends who brought normalcy to a less than normal environment. These guys, and the folks at CABS, made this journey much more enjoyable. Without them, I might have been teaching tennis in South Texas instead of expanding my career.

Finally, I thank my role models: my parents who are the

best, my grandfather Herman Winkle who is the most honest man I've ever met (and MaMaw for always making me laugh), Scott who embodies vision and leadership, and Coaches Acker and Peterson who define strength.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGMENTS.....	iv
<u>chapter</u>	<u>page</u>
INTRODUCTION.....	1
Need for Self-Appraisals.....	1
Weaknesses of Self-Appraisals.....	3
Impression Management.....	5
Self-Deception.....	7
Overview.....	9
Hypotheses.....	17
METHOD.....	17
Pilot Study.....	17
Participants.....	17
Procedure.....	18
Experiment Phase.....	19
Participants.....	19
Design.....	19
Procedure.....	19
Independent Variables.....	21
Impression Management Context.....	21
Self-Monitoring.....	22
Individual Differences Measures.....	22
Self-Monitoring Scale.....	22

<u>chapter</u>	<u>page</u>
BIDR Scale.....	23
Dependent Variables.....	25
Accuracy.....	25
Leniency.....	25
RESULTS.....	26
Correlation of Scales.....	30
DISCUSSION.....	31
REFERENCES.....	37
APPENDICES.....	43
A. INFORMED CONSENT FORM.....	44
B. SELF-MONITORING SCALE.....	46
C. SELF-DECEPTION SUBSCALE OF BIDR.....	48
D. IMPRESSION MANAGEMENT SUBSCALE OF BIDR....	50
E. SELF-RATING QUESTIONNAIRE.....	52
F. DEBRIEFING STATEMENT.....	54
G. SUMMARY TABLES FOR DEPENDENT VARIABLES....	56
VITA.....	75

Examining the Impact of Impression Management Context and Self-Monitoring on the Leniency and Accuracy of Self-Appraisals

Within the performance appraisal domain, self-appraisals of performance are increasingly investigated in research settings and utilized in applied settings. Self-ratings of performance commonly are used to supplement feedback from supervisors, impact contract negotiations, determine merit pay raises, and influence the termination of employment (Ashford, 1989). Despite the growing popularity of self-appraisals, past research has demonstrated that self-appraisals of performance are consistently less accurate than peer and supervisor ratings as well as objective measures of performance (Mabe and West, 1982). Enhancing self-appraisal utility in applied settings will likely result from studies which increase our understanding of the underlying factors which contribute to inaccurate and lenient self-ratings.

Need for Self-Appraisals

The growing need for accurate self-appraisals stems from dramatic changes in the structure and complexity of organizational settings. Specifically, team-oriented work environments, Quality of Work Life programs, and increased boundary spanning roles advocating self-development and

autonomy will lead to more frequent implementation of self-appraisals in applied settings (Sundstrom, Demeuse, & Futrell, 1990). With this in mind, effective self-assessment procedures may soon be required for organizations to prosper in increasingly dynamic and turbulent environments (Ashford, 1989).

In general, self-appraisals play a critical role in self-regulation, provide an easily accessible and trustworthy source of feedback, and are particularly valuable when environmental cues are ambiguous (Ashford, 1989). In addition, self-ratings improve communication between supervisors and subordinates, promote personal development, and clarify differences of opinion between supervisors and other managers (Harris and Schaubroeck, 1988). Clearly, this supplemental information strengthens the entire appraisal process in organizational settings (Klimoski and London, 1974).

Moreover, self-appraisal utilization may decrease an individual's defensiveness in the appraisal interview and increase his/her achievement motivation (Cascio, 1991). Also, self-appraisals may be more efficient, economical, and effective than traditional oral and written examinations within the hiring process (Anderson, Warner, and Spencer, 1984). Along these lines, self-appraisals may be a better evaluation tool than other conventional personnel selection

methods because they have high fairness, acceptable validity, and are relatively inexpensive (Muchinsky, 1986).

Finally, enhanced organizational fairness and justice are associated with self-appraisal utilization. Specifically, ratings of procedural fairness are significantly higher when ratees provide input into the appraisal process than when this opportunity is not provided (Bies & Shapiro, 1988). In sum, self-appraisals facilitate employee development (Campbell & Lee, 1988), increase the acceptance of decisions based on appraisal information (Farh, Werbel, & Bedeian, 1988), and enhance the communication between supervisors and subordinates regarding job content, performance criteria, and mutual expectations (Cardy & Dobbins, 1994).

Weaknesses of Self-Appraisals

Although support for self-appraisal utilization is widespread, many researchers are pessimistic about their practical utility. Clearly, inaccuracy and leniency pervade self-ratings of performance. Inaccuracy reflects judgment standards which are unduly positive or negative in the rating process. Leniency errors reflect inappropriate standards for judgment in which ratings are excessively positive (Bernardin & Beatty, 1984). The inability of individuals to objectively and reliably evaluate themselves leads to inaccurate, self-enhanced appraisals which can

seriously compromise their usefulness in organizations (Farh, Werbel & Bedeian, 1988).

With this in mind, Harris and Schaubroeck (1988) conducted a meta-analysis examining self, peer, and supervisor ratings and found that self-evaluations were approximately one-half standard deviation higher than supervisor ratings. In addition, a meta-analysis comparing self-rated ability and objective performance measures yielded only moderate correlations (Mabe & West, 1982). Similarly, DeNisi and Shaw (1977) found that correlations between self-rated ability and observed test performance were too small to have any practical significance. In fact, the widespread leniency and inaccuracy associated with self-appraisals has led some researchers to discount their utility as an evaluation tool and suggest the abandonment of self-appraisals entirely (Williams & Levy, 1992).

However, understanding the underlying factors which contribute to leniency effects may enhance their utility as an evaluation tool. Traditionally, researchers have argued that leniency is an intentional distortion of one's performance (Bernardin & Beatty, 1984). Consequently, leniency may be caused by a desire to be liked, a fear that coworkers will exaggerate their own ratings, and low standards for outstanding performance (Saal & Knight, 1988). However, recent research suggests that leniency may not be

limited to intentional motivational biases occurring at the time of the appraisal process (Hauenstein, 1992).

Unintentional biases may also contaminate self-appraisals. Broadly, then, potential factors underlying leniency effects will be examined.

Impression Management and Leniency in Self-Appraisals

Impression management is defined as the process by which people actively control and regulate the impressions that others form of them (Leary and Kowalski, 1990). In organizational settings, it is necessary to examine the goal-relevance of impression management techniques within the self-appraisal context. Clearly, self-ratings of performance directly impact how employees will be perceived by their superiors. Moreover, self-evaluations are often used in job selection, merit pay decisions, and job promotion considerations (Ashford, 1989). As a result, individuals will likely be motivated to appear confident and competent in self-rated performance in order to maximize the possibility of attaining work-related goals. Negative self-disclosures in self-reports may be construed by the employee as compromising his/her chances of maintaining a current position or moving up within the organizational ranks.

With this in mind, Jones (1990) discusses impression management processes in terms of environmental affordances, which he terms the ecological viewpoint. The ecological

viewpoint holds that individuals perceive dynamic stimulus information from events in their environment and alter their goal-directed behavior accordingly. More specifically, behavior is guided by perceptions of environmental affordances. This leads individuals to seek out information to determine how others can reinforce or punish them.

Normally, self-appraisals are evaluated by superiors who have power over the rater. This power asymmetry may result in exclusively positive self-disclosures in which negative information about the self is omitted (McArthur & Brown, 1983). Clearly, self-evaluations afford workers a unique opportunity from their environment to present themselves, to those who have power over them, as confident and competent workers.

Along these lines, Goffman's dramaturgical approach is also relevant in the discussion of impression management. This approach holds that life is analogous to a theatrical stage in which individuals are motivated to establish a desired tone within numerous, complex social interaction (Leary and Kowalski, 1990). Specifically, social interactions are viewed as experiences in which individuals enact certain roles in order to accomplish given motives. Thus, by assuming roles and acting them out, the worker attempts to navigate through often perilous social interactions to attain his/her underlying goals. So, in

constructing our social reality, we as "actors" work diligently to shape others' perceptions of us in order to obtain these goals. Within the context of self-appraisals, the worker's "mask" takes the form of self-evaluations in which the individual will likely present himself/herself to others in an excessively positive manner, resulting in inaccurate and lenient self-ratings.

In sum, impression management needs to be examined in self-appraisal research. Clearly, impression management processes explain part of the inaccuracy and leniency typically found in self-evaluations of performance in organizational settings.

Self-Deception and Leniency in Self-Appraisals

Self-deception may also contribute to lenient and inaccurate self-appraisals. Research indicates that people have inflated and exaggerated beliefs in their abilities (Meyer, 1980). For instance, most individuals tend to rate themselves significantly higher than others even when self-ratings are not disclosed to their supervisors. In confidential self-ratings, Meyer (1980) found, "The average self-appraisal of the 92 subjects who participated in the study was at the 78th percentile. That is, on average, each participant felt that he/she was performing better than about three-fourths of the other employees in similar jobs" (p. 292). Thus, although subjects were not motivated to

convey a positive self-image to a supervisor, who did not have access to the self-evaluation, they were still overly optimistic about their abilities and job performance.

In addition, individuals have unrealistically positive views of themselves, exaggerated beliefs in control over their environment, and unrealistic optimism (Taylor & Brown, 1988). These illusions are maintained because social and cognitive filters tend to highlight positive information and minimize negative information about the self. This phenomenon is due, in part, to people's general hesitancy to offer negative feedback to others. Furthermore, when negative information does penetrate these filters, its effects are usually temporary. Even recurrent negative information that does not get filtered out or forgotten is often relegated to pockets of incompetence in which the perceived importance of these weaknesses is minimized (Taylor & Brown, 1988).

Further, self-deception may be maintained because individuals tend to implicitly signal how they wish to be perceived by others and then surround themselves with individuals who respond to those cues in a manner consistent with the individual's self-opinion. Generally, these illusions serve various cognitive, affective, and social functions which benefit the individual (Taylor & Brown, 1988). They do not, however, enhance the utility of self-

appraisals and likely contribute to inaccurate and inflated evaluations of one's performance.

Finally, attribution theory also provides an explanation for self-deception processes in self-appraisals. This theory maintains that individuals attribute their own poor performance to proximal, situational factors, while attributing others' poor performance to inherent personality traits (Murphy & Cleveland, 1991). Conversely, individuals tend to view their own good performance as a reflection of distal, enduring personality characteristics, while attributing the good performance of others to situational factors. Thus, attribution theory may help explain why self-ratings are consistently higher than peer and supervisor ratings (Mabe & West, 1982).

Overview of Current Study

The current study examines the impact of impression management and self-deception on the leniency and inaccuracy associated with self-appraisals of performance. The distinction between these two processes is complex. Paulhus (1989) views self-deception and impression management as independent constructs and argues that self-deception occurs as a reaction to psychologically threatening phenomena that occur at the unconscious level. Support for this assertion comes from his studies in which self-deception was invariant across anonymous versus public disclosures of self-ratings,

whereas impression management varied significantly across conditions.

This finding lends tentative support for the assertion that self-deception and impression management have an additive effect on self-ratings, with self-deception being relatively constant. As Paulhus (1984) points out, "Taken together, these three studies argue strongly for a self-deception plus impression management theory of socially desirable responding in self-reports" (p. 606).

Nevertheless, the conceptualization of self-deception is more appropriate in terms of outcomes than in terms of constructs. In essence, Paulhus (1984) simply correlated items from numerous 'faking' scales and labeled the resulting factors as self-deception and impression management. However, these supposed constructs have not been thoroughly reviewed in the literature, like other constructs such as intelligence. This is important when one considers that on the self-deception subscale of the BIDR, if people have never questioned their sexuality, thought their parents hated them, or wanted to kill themselves to get back at someone they are, by definition, self-deceiving. It is difficult to make causal assertions simply based on test item categorization and labeling, particularly with the items used in the BIDR. Consequently, it is more appropriate to view the hypothesized baseline of leniency,

beyond true scores, as an outcome of self-deception caused by any number of constructs.

For instance, in negotiation exercises, high self-esteem has been shown to be related to inflated self-evaluations (Kramer, Newton, & Pommerenke, 1993). In addition, narcissism and defensive self-enhancement were related to differential levels of self-deception (Raskin, Novacek, & Hogan, 1991). Thus, individual difference variables such as self-esteem and narcissism likely impact self-deception leniency and are properly examined in terms of self-rating outcomes. Nevertheless, there is no reason to believe that self-deception varies across different environmental contexts.

Finally, although self-deception is viewed as an outcome, Paulhus's (1984) self-deception and impression management scales will be used in a supplemental fashion as evidence that people who display high levels of self-deception in the baseline of observed leniency should also score high on the self-deception questionnaire. Similarly, individuals who score high on the impression management scale will likely exhibit enhanced leniency beyond the baseline of self-deception when environmental affordances for leniency are available.

Overall, then, it is believed that self-deception initially inflates self-ratings beyond actual performance

levels. In addition, environmental factors which reward inflation (i.e. reward purposes of appraisal) may trigger impression management motives which increase leniency beyond self-deception. Finally, inflated self-ratings impacted by the impression management context are likely moderated by personality variables such as self-monitoring.

Impression Management Context

Unfortunately, it is difficult to determine conclusively if observed self-rating leniency is due to self-deception or impression management. However, there is no theoretical basis to support the assertion that self-deception varies as a function of the self-appraisal context. In other words, self-deception is believed to be relatively stable across self-appraisal contexts. This assertion is based on research such as Meyer's (1980) findings that in completely confidential self-ratings, the average self-appraised performance in comparison to others was at the 78th percentile. However, the degree of inflation and leniency in self-ratings has been shown to vary as a function of the impression management context.

For instance, purpose of appraisal (POA), categorized into reward (e.g., salary) and non-reward (e.g., feedback) conditions, has been demonstrated to differentially impact the degree of self-appraisal leniency (Zedeck & Cascio, 1982). Specifically, self-appraised performance in the

former condition is consistently more lenient and less accurate than in the latter condition (Bernardin & Orban, 1990; Farh & Werbel, 1986; Williams, DeNisi, Blencoe, & Cafferty, 1985). For example, students who were informed that borderline grades would be rounded upwards as a function of class participation displayed significantly higher self-ratings of participation than a control group of students who were told that grades would be unaffected by class participation (Farh & Werbel, 1986).

Similar results were found in an organizational setting in which self-ratings, used for promotional decisions in a law enforcement agency, were significantly more lenient than self-ratings used simply for feedback purposes (Bernardin & Orban, 1990). These findings are supported by research from social psychology that argues that people tend to inflate self-evaluations in order to attain positive outcomes (Ashford, 1989; Jones, 1990; McArthur & Brown, 1983; Leary & Kowalski, 1990; Mohrman & Lawler, 1983).

Given the theoretical assertion that self-deception is relatively constant across self-appraisal contexts, situational factors impacting leniency are believed to tap impression management processes. However, impression management contexts, such as POA, may not be sufficient to explain lenient self-appraisals beyond the baseline of inflation caused by self-deception. In other words, we must

examine other factors beyond environmental affordances to explain leniency effects associated with impression management. Given the diversity and complexity of personality differences, it is likely that specific personality variables such as self-monitoring may moderate the impression management context in explaining lenient self-ratings.

Self-Monitoring

Self-monitoring is a personality variable that has its roots in social psychology and has been widely applied in other settings (Snyder, 1986). Self-monitoring involves two information seeking processes: identifying appropriate behavior patterns for a given context and constructing appropriate behavior based on inner feelings, attitudes, and other personal attributes. High self-monitors are more motivated and better able to interpret social cues from their environment allowing them to control the images of themselves that they project in social interactions. Conversely, low self-monitors have much greater cross-situational consistency in behavior, are less concerned with social images, and act upon information from inner sources more than they do from environmental cues (Snyder, 1974).

In work related interactions, high self-monitors successfully interpret cues from the environment and alter their behavior in a goal directed and socially acceptable

manner to meet their job related goals. Not surprisingly, high self-monitors are better able than low self-monitors to correctly decode verbal expressions and accurately judge the intent of verbal responses (Mill, 1984). Furthermore, high self-monitors are more conformist and aware of external stimuli than low self-monitors (Haverkamp, 1994). Regarding leadership, high self-monitors are more likely to emerge as leaders because they are more sensitive to available social cues for appropriate behavior than are low self-monitors (Cronshaw & Ellis, 1991; Kent & Moss, 1990).

Accordingly, self-monitoring significantly correlates with average leader rankings and task relevant behavior (Zaccaro, Foti, and Kenny, 1991). This lends support for the assertion that trait based variance associated with emergent leadership is largely due to social perceptiveness and response flexibility associated with high self-monitors. Finally, high self-monitors show less consistency between reported attitudes and resultant behavior than low self-monitors (Dobbins, Farh, and Werbel, 1993).

Conversely, low self-monitors lack the ability or motivation to effectively employ self-presentation strategies and demonstrate significantly higher cross-situational consistency than high self-monitors. For instance, in terms of accurately reporting grade point averages, low self-monitors demonstrated significantly

higher congruence between rating behavior and attitudes across situations than did high self-monitors (Dobbins, Farh, and Werbel, 1993). Clearly, low self-monitors are directed by information from relevant inner sources rather than situational cues like their high self-monitoring counterparts. Therefore, performance ratings from low self-monitors, who rely primarily on information from biased, overly positive inner sources, will likely be more inflated than ratings from high self-monitors, who rely more on external cues from the environment than faulty inner sources.

Broadly speaking, the current study will examine self-deception and impression management processes in self-appraisals. Theoretically, self-deception is believed to occur regardless of the impression management context. However, impression management is believed to further inflate self-ratings beyond self-deception processes. As a result, subjects in contexts which are conducive to impression management processes should exhibit inflated self-ratings beyond subjects in conditions that are not conducive to impression management processes. However, responses to the impression management context are likely moderated by self-monitoring. Consequently, high SMS are expected to inflate their self-ratings significantly more than low SMS in conditions where impression management

affordances are salient. It is hoped that these contextual and personality variables will increase our understanding of self-deception and impression management processes which lead to lenient and inaccurate self-appraisals of performance.

Hypotheses

Hypothesis 1: Participants will overestimate their performance regardless of personality differences or impression management context.

Hypothesis 2: Participants will be less accurate and more lenient in conditions which provide environmental affordances for impression management processes.

Hypothesis 3: Self-Monitoring will moderate the impression management context such that high self-monitors will be less accurate and more lenient than low self-monitors in conditions which provide environmental affordances for impression management.

Method

Pilot Study

Participants

Participants were 15 undergraduate students from Virginia Polytechnic Institute and State University. Participants received one extra credit point for their participation in the pilot study. Subjects from the pilot

study were not eligible for participation in the experiment phase of the study.

Procedure

Two equivalent tasks were needed for the experiment phase of the study. Using Legos, subjects built as many planes, ships, and bridges as they could in three 15 minute intervals, for a total of 45 minutes. The tasks were counterbalanced to avoid practice and fatigue effects that might differentially advantage one task over the other in the absence of counterbalancing. The plane, ship, and bridge models consisted of ten, eleven, and ten pieces, respectively. A repeated measures ANOVA yielded an F value which was statistically significant ($p < .01$), reflecting significant differences in the omnibus test of the means. Two tailed paired comparison t-tests were then utilized. The mean number of correctly made planes ($M=22.8$, $SD=3.31$) and the mean number of correctly made ships ($M=23.7$, $SD=3.9$) were not significantly different ($t(14)=-.84$; $p > .05$). However, the average number of correctly made bridges ($M=16.67$, $SD=3.04$) was significantly lower than both the plane task ($t(14)=3.93$; $p < .05$) and the ship task ($t(14)=5.8$; $p < .05$). Consequently, the plane and ship models were retained for the experiment phase of the study.

Experiment Phase

Participants

Participants were 60 undergraduate students from Virginia Polytechnic Institute and State University. Participants received two extra credit points applied toward an undergraduate psychology class for their participation in the study.

Design. A 2 (male, female) X 2 (dimension, actual versus self-rating score) design was utilized to test group differences between actual and self-rated task performance in the non-reward POA. In addition, a 2 (male, female) X 2 (high SM, low SM) X 2 (no reward POA, reward POA) design was used to determine the impact of self-monitoring on the purpose of appraisal conditions. Because a similar number of males (N=28) and females (N=32) participated in the study, sex effects were investigated and included in the original run of all analyses.

Procedure

Upon entering the lab, participants were informed that they were involved in an experiment examining motor task ability involving a ship and boat building task. Participants were tested individually. Participants were then asked to complete a consent form and three questionnaires (see Appendices A, B, C, and D). The experimenter left the lab while the participant completed

the questionnaires. The experimenter then returned to the lab and placed a box of Legos and a model Lego plane (ship) on the table. Participants were read the following script, "You are now asked to build as many planes (ships) as you can in fifteen minutes. All planes (ships) must correctly and completely resemble the model plane (ship) that is provided, including the colors of the legos." The participants placed the models in a covered box which kept them from seeing how many models they built.

Fifteen minutes later, the experimenter returned to the lab and handed the participant the self-rating form for performance (see Appendix E) and the questionnaires (see Appendices C and D) and exited the lab. Once the participant completed the questionnaires, the experimenter returned to the lab with the model ship (plane) and read the previous task instructions with the following addition, "The top 5% of all performers on this task will receive an additional extra credit point." In actuality, all participants in the study received two extra credit points for their participation in the study, regardless of their performance. In terms of task counterbalancing, the first 30 participants completed the plane first (non-reward POA) and the ship second (reward POA). The remaining 30 subjects completed the ship first (non-reward POA) and the plane (reward POA) second.

Following the second task, the experimenter returned to the lab and handed the participant the self-rating questionnaire and debriefing form (see Appendix F). These forms were then collected and the experimenter explained the nature of the study to the participant and answered all questions. For each subject, the experiment lasted approximately forty five minutes.

Independent Variables

Impression Management Context. Impression management context reflects the process in which individuals tailor self-evaluations of performance in accordance with environmental affordances. In this context, purpose of appraisal, including reward and non-reward POA, impacts the leniency and accuracy of self-appraised performance.

Participants first completed the task in a non-reward POA condition. They were expected to be inaccurate and lenient because of genuinely being unable to rate themselves accurately, rather than being influenced by situational factors which promote impression management tactics. Participants then performed the task in a reward POA condition. They were told that the top five percent of all performers would receive an additional extra-credit point, although all subjects received two extra credit points. It was expected that impression management motives would be triggered in the reward POA condition, yielding more lenient

and inaccurate self-ratings beyond the self-deception processes associated with the non-reward condition.

Self-Monitoring. Because impression management conditions are not salient in the non-reward POA, both SM groups should be similarly lenient and inaccurate in the non-reward condition. Unlike low self-monitors, high self-monitors are expected to be significantly more lenient and inaccurate in the reward POA than in the non-reward POA. In addition, high self-monitors are expected to be significantly more lenient and inaccurate than the low SMS in the reward POA.

Individual Difference Measures

Self-Monitoring Scale. The first self-monitoring scale was developed by Snyder in 1974 who used 25 true-false questions to determine high versus low self-monitors. This questionnaire has been criticized because four of the five underlying factors that Snyder (1974) used to assess self-monitoring are positively related to social anxiety. This is incompatible with the effective social interactions characteristic of high self-monitors. For this reason, the Lennox and Wolfe (1984) 13 item Revised Self-Monitoring scale was utilized (Appendix B).

High versus low self-monitors were determined using a median split of overall self-monitoring scores. This questionnaire measures 'sensitivity to the expressive

behavior of others' (seven items) as well as the 'ability to modify self-presentation' (six items). The factor loadings for all items have ranged from .3 to .65 and coefficient alpha has been reported as .75 for this scale (Lennox and Wolfe, 1984). This measurement tool is preferred over Snyder's scale because it more narrowly defines self-monitoring, incorporates a Likert scale rather than true-false items, and contains items which are not positively associated with social anxiety.

Balanced Inventory of Desirable Responding (BIDR) Scale. Ten questions from the self-deception subscale of the BIDR and ten questions from the impression management subscale of the BIDR were utilized in the study. The items on the self-deception scale (Appendix C) were drawn from the self-deception questionnaire (items 1, 2, 5, 7 and 10) and the Edward's self-deception scale (items 3, 4, 6, 8 and 9) and involved items that are universally true but psychologically threatening in confronting parental and sexual conflicts. This scale is negatively keyed such that all 'No' responses, indicating self-deception, are summed for a total self-deception score.

In terms of the impression management questionnaire (Appendix D), items 1, 2, 6 and 7 are drawn from Paulhus' others' deception scale, items 3, 5 and 10 are gathered from the Marlow Crown scale, and items 4, 8 and 9 are drawn from

Wiggins' self-deception scale. These items reflect psychologically non-threatening and infrequently occurring items reflecting socially desirable behavior. The level of impression management is determined by summing all 'Yes' responses for items 1-3 and 6-8 and summing all 'No' responses on items 4-5 and 9-10. The aforementioned keyed responses reflect high levels of impression management.

Because it is impossible to unequivocally assert that self-deception occurs, the self-deception subscale of the BIDR was correlated with self-ratings in the non-reward purpose of appraisal. It was expected that the baseline of inflation predicted with self-ratings in non-reward purposes of appraisal would be correlated with scores from the self-deception scale. Also, further leniency in self-ratings predicted in the reward purpose of appraisal were expected to be correlated with scores from the impression management subscale of the BIDR. Finally, impression management varies across individuals so that the rank orders associated with the self-deception and impression management scales won't be identical. The BIDR scale was utilized to buttress the theoretical assertion that self-deception and impression management have an additive effect in terms of self-rating leniency.

Dependent Variables

It was predicted in the study that participants would generally overestimate their task performance and that their levels of overestimation would vary as a function of the purpose of appraisal and individual level of self-monitoring. These assertions were tested in terms of two dependent variables, accuracy and leniency.

Accuracy. Accuracy was defined as making an evaluation without error for a given standard and was objectively assessed in terms of the quantity and quality of models built. For each individual, the statistic for accuracy associated with ship building quantity was as follows: Quantity Accuracy = square root of $(O-T)^2$, where "O" denotes observed score (self-appraised score) and "T" denotes true (actual) score (see Appendix E, question 1). In addition, the accuracy statistic associated with ship building quality was as follows: Quality Accuracy = square root of $(O-T)^2$, where "O" denotes the observed score (self rated percentage of correctly made ships) and "T" denotes the true score (actual percentage of correctly made ships; see Appendix E, question 2). Mean observed and actual performance scores were also assessed for overall quantity and overall quality within the two purpose of appraisal conditions.

Leniency. Leniency was defined as an individual's tendency to unrealistically inflate self-ratings and was

assessed in terms of subjective comparisons with other participants. Leniency was assessed along the following four dimensions: quantity (see Appendix E, question 3) , quality (see Appendix E, question 4), effort (see Appendix E, questions 5-7), and conscientiousness (see Appendix E, questions 8-10). These dimensions were assessed in using a nine point Likert scale with the following anchors: one (worst), five (average), and nine (best).

Quantity was defined as the total estimated number models built. Quality was defined as the estimated percentage of correctly made models. Effort was defined as the degree to which an individual exerts mental and physical energy in a goal directed manner. Questions five, six, and seven were summed for each individual to attain his/her overall effort score. Conscientiousness was defined as the level of responsibility and ownership that the individual assumes in completing the task. Questions eight, nine, and ten were summed for each individual to attain his/her overall conscientiousness score.

Results

All hypotheses were tested using repeated measure ANOVAs. Because a similar number of males (N=28) and females (N=32) participated in the study, sex effects were also investigated and included in the original run of all

analyses. Descriptive statistics for all accuracy and leniency dimensions are provided in Tables 1 and 2.

The first hypothesis was that self-rated performance would be significantly inflated beyond actual performance regardless of purpose of appraisal conditions and self-monitoring. A 2 (male, female) X 2 (dimension, actual versus self-rating score) repeated measures ANOVA was used to test this hypothesis in terms of the overall quantity and quality of performance in the non-reward POA. The actual versus self-rating dimensions were only used for the first hypothesis because this hypothesis addressed overall differences between perceived and actual performance instead of differences in self-ratings as a function of the purpose of appraisal conditions. Because POA was not believed to impact self-ratings, differences between self-ratings and objective performance were only examined in the non-reward POA. Significant differences were found between self-rated and actual performance for both overall quantity and quality accuracy (Tables 3 and 4). However, results were in a direction opposite than predicted, with self-ratings being lower than true scores (Table 1). Also, the main effect for sex was significant with both the overall quantity and quality dimensions (Tables 3 and 4), with male self-ratings significantly higher than female self-ratings (Table 1). Overall, hypothesis one was rejected.

Hypothesis two stated that participants would be less accurate and more lenient when they were in the reward purpose of appraisal than when they were in the non-reward purpose of appraisal, reflecting impression management processes. The observed self-rating severity rendered the second hypothesis unrealistic in terms of accuracy. Inflation in the reward POA beyond the non-reward POA would actually enhance accuracy, given the initial baseline of severity, rather than increasing inaccuracy.

For the accuracy dimensions with the second hypothesis, two 2 (male, female) X 2 (no reward POA, reward POA) repeated measure ANOVAs were utilized. Difference scores varied significantly as a function of POA conditions for quantity accuracy but not quality accuracy (Tables 5 and 6). Thus, participants were significantly more accurate and less severe for quantity accuracy when they were in the reward POA condition. Sex effects and interactions were non-significant for the accuracy dimensions (Tables 5 and 6). Hypothesis two was rejected in terms of the accuracy dimensions, although the trend of decreased severity (given the initial baseline of severity) with quantity accuracy is consistent with the spirit of the second hypothesis.

For the leniency dimensions with the second hypothesis, four 2 (male, female) X 2 (no reward POA, reward POA) repeated measure ANOVAs were used to determine if self-

ratings were more lenient when participants were in the reward POA condition. For quantity and quality leniency, the main effects for sex were significant with men rating themselves higher than women (Tables 7 and 8), although the main effects for purpose of appraisal and the interaction terms were non-significant. Thus, self-ratings were not significantly more inflated when participants were in reward POA condition with these leniency dimensions. However, self-ratings were significantly more inflated with effort and conscientious leniency dimensions when participants were in the reward POA condition (Tables 9 and 10). Sex and interaction effects were non-significant for these variables. Thus, support for hypothesis two was provided with two of the four leniency dimensions.

The third hypothesis stated that self-monitoring would moderate self-ratings in the reward purpose of appraisal condition. To test this hypothesis, six 2 (male, female) X 2 (high self-monitors, low self-monitors) X 2 (no reward POA, reward POA) repeated measure ANOVAs were utilized. Main effects and interaction effects with self-monitoring were non-significant for all accuracy and leniency dimensions (Tables 11-16). For all other variables, the statistical significance of the F values in the full model was congruent with the partial model. In addition, an exploratory investigation of impression management was

conducted. Impression management replaced self-monitoring in the aforementioned design. Differences between high and low impression managers were non-significant and the overall results mirrored those found with self-monitoring. Broadly, hypothesis three was rejected.

Correlations of Scales

In order to bolster evidence that initial levels of inflation reflected self-deception in the non-reward POA and further inflation indicated impression management in the reward POA, the self-deception (Table 17) and impression management (Table 18) subscales from Paulhus were utilized using one-tailed significance tests. Correlations between self-deception scores and the dependent variables in the non-reward POA condition would provide indirect support that self-ratings in non-reward POA conditions are impacted by self-deception. In addition, correlations between impression management scores and the dependent variables in the reward POA condition would provide indirect support that self-ratings in reward POA conditions are impacted by impression management. Changes in rank order will result with the two scales because impression management varies across individuals.

Before examining the results of these correlations, it should be noted that the average impression management score was 4.68 ($SD=1.81$) and the average self-deception score was

7.32 ($SD=1.98$). Ceiling effects with the self-deception scale may have limited its utility in interpreting the results. The correlations for the self-deception scale are provided in Table 17 and the correlations for the impression management scale are provided in Table 18.

Correlations between the self-deception scale and all dependent variables were non-significant in the non-reward POA condition. Correlations between the impression management scale and the dependent variables were non-significant in the reward POA condition with the exception of the effort leniency dimension ($r=-.31$), which was negatively ($p<.05$) related to impression management in the reward POA condition (Table 18). Broadly, these correlations did not provide indirect support for the assertion that self-ratings in the non-reward POA condition were influenced by self-deception and that self-ratings in the reward POA condition were impacted by impression management.

Discussion

The goal of the current study was to investigate the underlying factors which contribute to lenient and inaccurate self-appraisals. Although past research has focused on specific variables which impact leniency and accuracy, there is often a lack of theoretical development in explaining *why* self-ratings are problematic.

Accordingly, this study set out to test the assertion that self-deception processes initially inflate self-ratings and that impression management processes, in certain contexts, further increase self-rating inflation. Surprisingly, an initial baseline of severity, not inflation, was found with self-rated performance. Before examining the theoretical implications of this severity, it is important to examine why overly severe ratings occurred in the first place.

The most obvious explanation for these results is that the task was simply too unfamiliar. Thus, participants did not have enough information to formulate a realistic idea of how many models that they built. Researchers have found that participants who are unfamiliar with appraisal procedures tend to underrate their performance in comparison to participants with more information about, and knowledge of, the appraisal process (Williams & Levy, 1992). Because the justification for using appraisals was hidden from the subjects to avoid contaminating effects and because they were unaware that self-appraisals would even occur until presented with self-rating forms, the unfamiliarity with the appraisal process may have lead to self-rating severity.

Finally, the severity effects may have been influenced by task simplicity. Researchers found, in calibration accuracy studies with dichotomous items, that participants

rated themselves higher on difficult tasks and more severely on easier tasks (Lichtenstein & Fischhoff, 1977). The task used in the current study was very simple, as illustrated by the extremely high percentage of accurately made models. Although generalizing from test responses to ship/plane building is tenuous, task simplicity may have contributed to the overly severe self-ratings. Perhaps task simplicity, in combination with task unfamiliarity, led to the overly severe self-ratings.

How, then, do the observed severity effects impact the theoretical framework of the current study? Broadly, participants likely had insufficient information to reasonably evaluate themselves because of a simplistic, overly unfamiliar task. Also, self-deception scales were not significantly correlated with the dependent variables in the non-reward purpose of appraisal. Thus, definitive conclusions about self-deception are impossible.

Nevertheless, participants rated their performance with the leniency dimensions very favorably in comparison to others (Table 2). Further, these comparative ratings were favorable despite self-ratings which were unduly severe in comparison to actual performance. More importantly, these favorable ratings occurred in the absence of salient environmental affordances for inflation (i.e., non-reward

purpose of appraisal). Still, the hypothesis concerning self-deception, as stated, was not supported.

In addition, the importance of impression management was illustrated. Self-ratings were significantly higher when participants were in the reward purpose of appraisal for both the effort and conscientious leniency dimensions (Tables 9 & 10). Further, self-ratings were significantly less severe when participants were in the reward purpose of appraisal condition for quantity accuracy (Table 5). Based on these results, impression management seemed to have partially impacted self-rated leniency and inaccuracy. Moreover, impression management processes should be more salient in applied settings than in the current study because environmental affordances for inflated self-ratings (e.g., salary raises, promotions) are likely more potent. Thus, partial support for the theoretical model, with self-deception and impression management, is provided.

However, the impact of self-monitoring in this model was not demonstrated. In fact, the POA X SM interaction terms seldom approached significance for any of the accuracy and leniency dimensions. Further investigations with self-monitoring may be well served to use extreme group designs to determine if strongly high or strongly low self-monitors react differently to the purpose of appraisal manipulation.

Finally, the impact of gender on self-rating accuracy and leniency should be examined in future self-appraisal research. In the current study, the main effects for gender were significant for half of the dimensions examined. With the exception of overall rated quality and conscientious leniency, men rated themselves substantially higher than females. Because a pro-male bias traditionally occurs with performance appraisals in applied settings (Dobbins, Cardy, & Truxillo, 1988) despite equivalent performance (Basow, 1992), future research should investigate the extent to which women may hinder their own professional progress through less aggrandizing self-evaluations. It is conceivable that the 'glass ceiling' effect, marked by limited opportunities for women beyond middle management, may result from less inflated self-evaluations by women in addition to a pro-male bias from supervisors. Specific hypotheses about female self-ratings, within the context of self-deception and impression management, would be useful.

In conclusion, mixed support for this theoretical model was provided. Although self-ratings were overly severe when participants were in the non-reward purpose of appraisal, subjects still rated themselves favorably in comparison to others in the non-reward POA condition. In addition, the influence of impression management processes were likely demonstrated by less severe (quantity accuracy) and more

inflated (effort and conscientious leniency) self-ratings in the reward, versus non-reward, purpose of appraisal. However, self-monitoring was a non-factor within this theoretical framework in the current study. Future research should utilize more familiar tasks to study self-deception, stronger reward manipulations for impression management, and extreme group designs with self-monitoring to further test the proposed model. Past research with self-appraisals has often consisted of 'fishing for variables' to decrease leniency without examining the underlying causes of such inflation. Future investigations of the proposed theoretical model may increase our understanding the factors which contribute to lenient and inaccurate self-ratings of performance.

References

Anderson, C. D., Warner, J. L., & Spencer, C. C. (1984). Inflation Bias in Self-Assessment Examinations: Implications for Valid Employee Selection. Journal of Applied Psychology, 69, 574-580.

Ashford, S. J. (1989). Self-Assessments in Organizations: A Literature Review and Integrative Model. Research in Organizational Behavior, 11, 133-174.

Basow, S. A. (1992). Gender Stereotypes and Roles. Pacific Grove, CA. Brooks/Cole Publishing Company.

Bernardin, H. J., & Beatty, R. W. (1984). Performance appraisal: Assessing human behavior at work, Boston: Kent.

Bernardin, H. J., & Orban, J. A. (1990). Leniency effect as a function of rating format, purpose of appraisal, and rater individual differences. Journal of Business and Psychology, 5, 197-211.

Bies, R. J., & Shapiro, D. L. (1988). Voice and justification: Their influence on procedural fairness judgments. Academy of Management Journal, 31, 676-685.

Campbell, D., & Lee, C. (1988). Self-appraisal in performance evaluation: Development vs. evaluation. Academy of Management Review, 13, 302-314.

Cardy, R. L., & Dobbins, G. H. (1994). Performance appraisal: Alternative Perspectives, Cincinnati, Ohio: Southwestern Publishing Co.

Cascio, W. F. (1991). Applied Psychology in Personnel Management, 4th ed, New Jersey: Simon and Shuster.

Cawley, B. D., & Levy, P. E. (1995). Performance rating discrepancies: a direct test of multiple explanations. Presented at the Tenth Annual Conference of the Society for Industrial Organizational Psychology, Inc. May 19-21. Orlando, Florida.

Cronshaw, S. F., & Ellis, R. J. (1991). A process investigation self-monitoring and leader emergence. Small Group Research, 22, 403-420.

DeNisi, A. S., & Shaw, J. B. (1977). Investigation of the uses of self-reports of abilities. Journal of Applied Psychology, 62, 641-644.

Dobbins, G. H., Cardy, R. L., & Truxillo, D. M. (1988). The effects of purpose of appraisal and individual differences in stereotypes of women on sex differences in performance ratings: A laboratory and field study. Journal of Applied Psychology, 73, 551-558.

Dobbins, G. H., Farh, J. L., & Werbel, J. D. (1993). The influence of self-monitoring on inflation of grade-point averages for research and selection processes. Journal of Applied Social Psychology, 23, 321-334.

Farh, J., Dobbins, G. H., & Cheng, B. (1991). Cultural relativity in action: a comparison of self-ratings made by chinese and U.S. workers. Personnel Psychology, 44, 129-145.

Farh, J., & Werbel, J. D. (1986). Effects of purpose of appraisal and expectation of validation on self-appraisal leniency. Journal of Applied Psychology, 71, 527-529.

Farh, J., Werbel, J. D., & Bedeian, A. G. (1988). An empirical investigation of self-appraisal-biased performance evaluation. Personnel Psychology, 41, 141-156.

Harris, M. M., & Schaubroeck, J. (1988). A Meta-Analysis of Self-Supervisor, Self-Peer, and Peer-Supervisor Ratings. Personnel Psychology, 41, 43-62.

Hauenstein, N. M. A. (1992). An information-processing approach to leniency in performance judgments. Journal of Applied Psychology, 77, 1-9.

Haverkamp, B. E. (1994). Using assessment in counseling supervision: Individual differences in self-monitoring. Measurement and Evaluation in Counseling and Development, 27, 316-324.

Jones, E. E. (1990). Interpersonal perception. New York: Freeman.

Kramer, R. M., Newton, E., & Pommerenke, P. L. (1993). Self-enhancement biases and negotiator judgment: Effects of self-esteem and mood. Organizational Behavior and Human Decision Processes, 56, 110-133.

Kent, R. L., & Moss, S. E. (1990). Self-monitoring as a predictor of leader emergence. Psychological Reports, 66, 875-881.

Klimoski, R.J., & London, M. (1974). Role of the rater in performance appraisal. Journal of Applied Psychology, 59, 445-451.

Leary, M. R., & Kowalski, R. M. (1990). The two components of impression management: A literature review and theoretical integration. Psychological Bulletin, 107, 34-47.

Lennox, R. D., & Wolfe, R. N. (1984). Revision of the self-monitoring scale. Journal of Personality and Social Psychology, 46, 1349-1364.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? Organizational Behavior and Human Performance, 20, 159-183.

Mabe III, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis, Journal of Applied Psychology, 67, 280-296.

McArthur, L. Z., & Baron, R. M. (1983). Toward an Ecological Theory of Social Perception. Psychological Review, 90, 215-238.

Meyer, H. H. (1980). Self-Appraisal of job performance. Personnel Psychology, 33, 291-295.

Mill, J. (1984). High and low self-monitoring individuals: Their decoding skills and empathic expression. Journal of Personality, 52, 372-388.

Mohrman, A. M., & Lawler, E. E. (1983). Motivation and performance appraisal behavior. In F. Landy, S. Zedek, & J.

Cleveland (Eds.), Performance measurement and theory, New Jersey: Erlbaum.

Muchinsky, P. M. (1986). Personnel selection methods. In C. L. Cooper & I. T. Robertson (Eds.), International Review of Industrial and Organizational Psychology, New York: John Wiley and Sons.

Murphy, K. R., & Cleveland, J. N. (1991). Performance Appraisal: An Organizational Perspective, Boston: Allyn and Bacon.

Paulhus, D. L. (1989). Socially desirable responding: Some new solutions to old problems. In D. M. Buss & N. Cantor (Eds.), Personality Psychology: Recent Trends and Emerging Directions. New York: Springer-Verlag.

Raskin, J., Novacek, J., & Hogan, R. (1991). Narcissism, self-esteem, and defensiveness self-enhancement. Journal of Personality, 59, 19-38.

Saal, F. E., & Knight, P. A. (1988). Industrial/Organizational Psychology: Science and Practice, Pacific Grove, California: Brooks/Cole.

Snyder, M. (1974). Self-monitoring of expressive behavior. Journal of Personality and Social Psychology, 30, 526-537.

Snyder, M. (1986). Public Appearances/Private Reality, New York: W. H. Freeman and Company.

Sundstrom, E., Demeuse, K. P., & Futrell, D. (1990). Work teams: Application and effectiveness. American Psychologist, 45, 120-133.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. Psychological Bulletin, 103, 193-210.

Williams, J. R., & Levy, P. E. (1992). The effects of perceived system knowledge on the agreement between self-ratings and supervisor ratings. Personnel Psychology, 45, 835-846.

Williams, K. T., DeNisi, A. S., Blencoe, A. G., & Cafferty, T. P. (1985). The role of appraisal purpose: Effects of purpose on information acquisition and utilization. Organizational Behavior and Human Decision Processes, 35, 314-339.

Zaccaro, S. J., Foti, R. J., & Kenny, D. A. (1991). Self-monitoring and trait-based variance in leadership: An investigation of leader flexibility across multiple group situations. Journal of Applied Psychology, 76, 308-315.

Zedeck, S., & Cascio, W. F. (1982). Performance appraisal decisions as a function of rater training and purpose of appraisal. Journal of Applied Psychology, 67, 752-758.

APPENDICES

APPENDIX A:
INFORMED CONSENT FORM

TITLE OF EXPERIMENT: Motor Skill Abilities in College Students.
EXPERIMENT #

1. PURPOSE OF EXPERIMENT

You are invited to participate in a study about motor skill abilities in college students.

2. PROCEDURE TO BE FOLLOWED IN THE STUDY

To accomplish the goals of the study, you will be asked to build model ships for 15 minutes.

3. CONFIDENTIALITY OF RESULTS

Results from the study will be completely confidential.

4. DISCOMFORTS AND RISKS FROM PARTICIPATING IN THE STUDY

There are no risks or discomforts associated with this study.

5. EXPECTED BENEFITS

This study hopes to further the understanding of factors involving motor skill performance.

6. FREEDOM TO WITHDRAW

You are free to withdraw from participation in this study at any time without penalty.

7. EXTRA CREDIT

For participation in this study, you will receive 1 extra credit point toward the introductory psychology class or any other undergraduate class with instructor approval.

8. USE OF RESEARCH DATA

The information from this research may be used for scientific or educational purposes. It may be presented at scientific meetings and/or published and reproduced in professional journals and books, or used for any other purpose that Virginia Tech's Department of Psychology considers proper in the interest of education, knowledge, or research.

9. APPROVAL OF RESEARCH

This research project has been approved by the Human Subjects Committee of the Department of Psychology and by the Institutional Review Board of Virginia Tech.

10. SUBJECT'S PERMISSION

I have read and understand the above description of the study. I have had an opportunity to ask questions and have had them all answered. I hereby acknowledge the above and give my voluntary consent for participation in this study. I further understand that if I participate this study, I may withdraw at any time without penalty. I understand that if I have any questions regarding this research and its conduct, I should contact any of the following persons below:

Primary researcher: Josh Williams	961-3221
Faculty advisor: Dr. Neil M. Hauenstein	231-5716
Chair, IRB: Ernest Stout	231-9359
Chairperson, HSC: RM Eisler	231-7001

Participant's signature _____

Date: _____ Participant's ID _____

APPENDIX B:
SELF-MONITORING SCALE

Lennox and Wolfe (1984) Revised Self-Monitoring Scale:

Rate how well each item relates to you using the following format: 0) always false, 1) generally false, 2) somewhat false, 3) somewhat true, 4) generally true, 5) always true.

1. In social situations, I have the ability to alter my behavior if I feel that something else is called for.
2. I have the ability to control the way I come across to people, depending on the impression I wish to give them.
3. When I feel that the image I am portraying isn't working, I can readily change it to something that does.
4. I have trouble changing my behavior to suit different people and different situations.*
5. I have found that I can adjust my behavior to meet the requirements of any situation I find myself in.
6. Even when it might be to my advantage, I have difficulty putting up a good front.*
7. Once I know what the situation calls for, it's easy for me to regulate my actions accordingly.
8. I am often able to read people's true emotions correctly through their eyes.
9. In conversations, I am sensitive to even the slightest change in the facial expression of the person I'm conversing with.
10. My powers of intuition are quite good when it comes to understanding others' emotions and motives.
11. I can usually tell when others consider a joke to be in bad taste, even though they may laugh convincingly.
12. I can usually tell when I've said something inappropriate by reading it in the listener's eyes.
13. If someone is lying to me, I usually know it at once from that person's manner of expression.

*denotes reversed scoring

APPENDIX C:
SELF-DECEPTION SUBSCALE OF BIDR

SCALE ONE:

For the following items, please indicate yes or no.

1. Have you ever enjoyed your bowel movements?
2. Have you ever been uncertain as to whether or not you are homosexual?
3. People often disappoint me.
4. Life is a strain for me most of the time.
5. Have you ever doubted your sexual adequacy.
6. I worry quite a bit over possible misfortunes.
7. Have you ever thought that your parents hated you?
8. I have several times given up doing something because I thought too little of my ability.
9. In a group of people I have trouble thinking of the right things to talk about.
10. Have you ever thought of committing suicide in order to get back at someone?

APPENDIX D:
IMPRESSION MANAGEMENT SUBSCALE OF BIDR

SCALE TWO:

For the following items, please indicate yes or no.

1. Do you tell the truth?
2. When you take sick-leave from work or school, are you as sick as you say you are?
3. I am always courteous, even to people who are disagreeable.
4. Once in a while, I laugh at dirty jokes.
5. I sometimes try to get even, rather than forgive and forget.
6. I always apologize to others for my mistakes.
7. Would you declare everything at customs, even if you knew that you could never be found out?
8. I never attend a sexy show if I can avoid it.
9. Sometimes at elections I vote for candidates I know little about.
10. I am sometimes irritated by people who ask favors of me.

APPENDIX E:
SELF-RATING QUESTIONNAIRE

1. How many planes (ships) did you build?
2. What percentage of the planes (ships) that you built exactly resembled the model that was provided for you?

For the remaining questions, rate yourself in comparison to other undergraduate psychology students on this task from 1 to 9 using the following scale: 1) worst, 5) average, 9) best.

3. Total number of planes (ships) built.
4. Quality of the planes (ships) that you built (i.e. your ships exactly resembled the model).
5. For this task, my level of effort was high.
6. I tried extremely hard on this task.
7. For this task, I was determined to perform at a high level.
8. For this task, I performed as well as I could.
9. For this task, it was important that I perform at a high level.
10. It was my responsibility to do the best job that I could on this task.

APPENDIX F:
DEBRIEFING STATEMENT

The study you just participated in did examine motor skill performance as reflected in the ship building task. However, we also examined the effects of the situation and personality (i.e. self-monitoring) on self-rated performance. Therefore, the quality and quantity of the number of planes and ships you built will be assessed in terms of whether or not you were in a situation that would reward positive self-ratings as well as the scores you receive from the self-monitoring scale. Briefly that scale is used to assess the degree to which you are motivated and able to assess cues from your environment and alter your behavior accordingly. Individuals who have high scores on this questionnaire and who are in conditions which reward inflated self-ratings are expected to have self-appraisals that are most lenient and inaccurate. Also, high scores on the self-deception questionnaire are expected to relate to high levels of inflation. Similarly, high scores on the impression management scale are expected to be related to inflated self-ratings in the reward condition (second task).

In the reward conditions, subjects were told that they would receive an additional extra credit point if they were one of the top 5% of all performers. In actuality, all participants receive two extra credit points, regardless of their performance, for their participation in the study. Participants are encouraged to contact the primary researcher upon completion of the study to determine if extra credit points were obtained.

If you would like further information on this study or would like a copy of the results, please contact Josh Williams at 961-3221 or Dr. Neil Hauenstein at 231-5716. Thank you for your participation.

Participant's signature _____

Date _____

Participant's ID _____

APPENDIX G:
SUMMARY TABLES FOR DEPENDENT VARIABLES

Table 1

Means and Standard Deviations for Accuracy Dimensions

Dimension	Purpose of Appraisal Condition	
	Non-reward	Reward
Rated Quantity		
Overall	19.82 (8.52)	21.88 (8.64)
Male	23.11 (8.85)	24.04 (8.98)
Female	16.94 (7.18)	20 (8)
High Self-Monitors	22.24 (9.36)	24.3 (8.24)
Low Self-Monitors	16.85 (6.35)	18.93 (8.32)
Actual Quantity		
Overall	25.6 (7.61)	26.13 (8.31)
Male	27.36 (7.54)	28.61 (9.15)
Female	24.06 (7.61)	23.97 (6.94)
High Self-Monitors	27.58 (7.8)	28 (7.15)
Low Self-Monitors	23.19 (6.74)	23.85 (9.16)
Rated Quality		
Overall	.9735 (.06)	.979 (.04)
Male	.9625 (.07)	.9679 (.05)
Female	.9831 (.04)	.9888 (.03)
High Self-Monitors	.967 (.07)	.9788 (.04)
Low Self-Monitors	.9815 (.04)	.9793 (.04)
Actual Quality		
Overall	.9897 (.04)	.9963 (.01)
Male	.9789 (.06)	.9954 (.02)
Female	.9991 (.01)	.9972 (.01)
High Self-Monitors	.9821 (.06)	.9976 (.01)
Low Self-Monitors	.9989 (.01)	.9948 (.02)
Quantity Accuracy		
Overall	8.55 (8.02)	5.25 (5.8)
Male	6.89 (5.31)	5.64 (6.18)
Female	10 (9.65)	4.91 (5.53)
High Self-Monitors	7.7 (5.53)	5.09 (6.02)
Low Self-Monitors	9.59 (10.31)	5.44 (5.63)
Quality Accuracy		
Overall	.0285 (.06)	.0218 (.04)
Male	.0304 (.05)	.0239 (.05)
Female	.0209 (.06)	.02 (.03)
High Self-Monitors	.0367 (.07)	.0203 (.04)
Low Self-Monitors	.0185 (.04)	.0237 (.04)

Note. Standard Deviations are in parentheses.

Table 2

Means and Standard Deviations for Leniency Dimensions

<u>Dimension</u>	<u>Purpose of Appraisal Condition</u>	
	Non-reward	Reward
Quantity Leniency		
Overall	6.23 (1.16)	6.38 (1.24)
Male	6.82 (1.06)	6.93 (1.25)
Female	5.72 (.99)	5.91 (1.03)
High Self-Monitors	6.52 (1.12)	6.67 (1.16)
Low Self-Monitors	5.89 (1.12)	6.04 (1.26)
Quality Leniency		
Overall	6.85 (1.56)	6.82 (1.65)
Male	7.39 (1.45)	7.61 (1.57)
Female	6.38 (1.52)	6.13 (1.41)
High Self-Monitors	7.27 (1.4)	7.15 (1.54)
Low Self-Monitors	6.33 (1.62)	6.41 (1.72)
Effort Leniency		
Overall	20.7 (5.22)	22.65 (4.3)
Male	21.61 (5.12)	23.89 (3.41)
Female	19.91 (5.25)	21.56 (4.73)
High Self-Monitors	21.64 (3.99)	22.91 (3.5)
Low Self-Monitors	19.56 (6.3)	22.33 (5.16)
Conscientious Leniency		
Overall	21.13 (5.06)	23.28 (3.47)
Male	21.04 (5.76)	23.21 (3.36)
Female	21.22 (4.46)	23.34 (3.62)
High Self-Monitors	22 (4.23)	23.82 (3.24)
Low Self-Monitors	20.07 (5.83)	22.63 (3.69)

Note. Standard Deviations are in parentheses.

Table 3

Anova Table for Partial Model with Rated versus True Score
Quantity in the non-reward POA

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	728.1	8.89**
Residual	58	84.24	
<u>Within Subjects</u>			
Dimension	1	1081.13	33.71**
Dimension X Gender	1	14.11	.42
Residual	58	13.19	

*denotes $p < .05$

**denotes $p < .01$

Table 4

Anova Table for Partial Model with Rated versus True Score Quality in the non-reward POA

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	.02	5.23**
Residual	58	.01	
<u>Within Subjects</u>			
Dimension	1	.02	8.85**
Dimension X Gender	1	.01	.59
Residual	58	.01	

*denotes $p < .05$

**denotes $p < .01$

Table 5

Anova Table for Partial Model with Quantity Accuracy

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	41.96	.72
Residual	58	58.39	
<u>Within Subjects</u>			
Purpose of Appraisal	1	300.48	7.77**
Gender X Purpose of Appraisal	1	110.32	2.85
Residual	58	38.67	

*denotes $p < .05$

**denotes $p < .01$

Table 6

Anova Table for Partial Model with Quality Accuracy

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	.01	.12
Residual	58	.01	
<u>Within Subjects</u>			
Purpose of Appraisal	1	.01	.78
Gender X Purpose of Appraisal	1	.01	.01
Residual	58	.01	

*denotes $p < .05$

**denotes $p < .01$

Table 7

Anova Table for Partial Model with Quantity Leniency

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	33.72	18.21**
Residual	58	1.85	
<u>Within Subjects</u>			
Purpose of Appraisal	1	.65	1.35
Gender X Purpose of Appraisal	1	.05	.75
Residual	58	.48	

*denotes $p < .05$

**denotes $p < .01$

Table 8

Anova Table for Partial Model with Quality Leniency

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	46.67	12.83**
Residual	58	3.64	
<u>Within Subjects</u>			
Purpose of Appraisal	1	.01	.01
Gender X Purpose of Appraisal	1	1.61	2.06
Residual	58	.78	

*denotes $p < .05$

**denotes $p < .01$

Table 9

Anova Table for Partial Model with Effort Leniency

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	121.34	3.21
Residual	58	37.84	
<u>Within Subjects</u>			
Purpose of Appraisal	1	116.03	17.83**
Gender X Purpose of Appraisal	1	2.96	.45
Residual	58	6.51	

*denotes $p < .05$

**denotes $p < .01$

Table 10

Anova Table for Partial Model with Conscientious Leniency

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	.73	.02
Residual	58	30.53	
<u>Within Subjects</u>			
Purpose of Appraisal	1	138.29	17.83**
Gender X Purpose of Appraisal	1	.02	.01
Residual	58	7.76	

*denotes $p < .05$

**denotes $p < .01$

Table 11

Anova Table for Full Model with Quantity Accuracy

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	24.82	.42
Self-Monitoring	1	29.45	.49
Gender X Self-Monitoring	1	12.63	.21
Residual	56	59.79	
<u>Within Subjects</u>			
Purpose of Appraisal	1	255.54	6.52*
Gender X Purpose of Appraisal	1	114.07	2.91
Self-Monitoring X Purpose of Appraisal	1	2.89	.07
Gender X Self-Monitoring X Purpose of Appraisal	1	41.88	1.07
Residual	56	39.2	

*denotes $p < .05$

**denotes $p < .01$

Table 12

Anova Table for Full Model with Quality Accuracy

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	.01	.09
Self-Monitoring	1	.01	.33
Gender X Self-Monitoring	1	.01	.31
Residual	56	.01	
<u>Within Subjects</u>			
Purpose of Appraisal	1	.01	.49
Gender X Purpose of Appraisal	1	.01	.08
Self-Monitoring X Purpose of Appraisal	1	.01	2.09
Gender X Self-Monitoring X Purpose of Appraisal	1	.01	.01
Residual	56	.01	

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

Table 13

Anova Table for Full Model with Quantity Leniency

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	28.74	15.99**
Self-Monitoring	1	5.46	3.04
Gender X Self-Monitoring	1	.78	.44
Residual	56	1.8	
<u>Within Subjects</u>			
Purpose of Appraisal	1	.33	.7
Gender X Purpose of Appraisal	1	.13	.28
Self-Monitoring X Purpose of Appraisal	1	.03	.06
Gender X Self-Monitoring X Purpose of Appraisal	1	1.23	2.59
Residual	56	.47	

*denotes $p < .05$

**denotes $p < .01$

Table 14

Anova Table for Full Model with Quality Leniency

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	37.62	10.59**
Self-Monitoring	1	11.41	3.21
Gender X Self-Monitoring	1	.22	.06
Residual	56	3.55	
<u>Within Subjects</u>			
Purpose of Appraisal	1	.05	.06
Gender X Purpose of Appraisal	1	1.5	1.93
Self-Monitoring X Purpose of Appraisal	1	.41	.53
Gender X Self-Monitoring X Purpose of Appraisal	1	1.25	1.6
Residual	56	.78	

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

Table 15

Anova Table for Full Model with Effort Leniency

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	102.19	2.65
Self-Monitoring	1	26.31	.68
Gender X Self-Monitoring	1	4.95	.13
Residual	56	38.58	
<u>Within Subjects</u>			
Purpose of Appraisal	1	128.16	20.23**
Gender X Purpose of Appraisal	1	7.13	1.12
Self-Monitoring X Purpose of Appraisal	1	21.36	3.37
Gender X Self-Monitoring X Purpose of Appraisal	1	2.76	.44

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

Table 16

Anova Table for Full Model with Conscientious Leniency

Source	df	MS	F
<u>Between Subjects</u>			
Gender	1	11.24	.38
Self-Monitoring	1	91.04	3.12
Gender X Self-Monitoring	1	58.97	2.02
Residual	56	29.19	
<u>Within Subjects</u>			
Purpose of Appraisal	1	144.17	18.23**
Gender X Purpose of Appraisal	1	.5	.06
Self-Monitoring X Purpose of Appraisal	1	4.95	.63
Gender X Self-Monitoring X Purpose of Appraisal	1	2.55	.32
Residual	56	7.91	

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

Table 17

Correlations Between Self-Deception Scale and Dependent Variables in the non-reward POA

Quantity Accuracy	.13
Quality Accuracy	.21
Quantity Leniency	.06
Quality Leniency	-.06
Effort Leniency	-.11
Conscientious Leniency	-.04

Note. All correlations are non-significant at the $p < .05$ level.

Table 18

Correlations Between Impression Management Scale and
Dependent Variables in the reward POA

Quantity Accuracy	-.15
Quality Accuracy	.08
Quantity Leniency	.05
Quality Leniency	-.11
Effort Leniency	-.31*
Conscientious Leniency	-.04

*Denotes statistically significant at $p < .05$ level.

VITA

JOSHUA H. WILLIAMS

PERSONAL INFORMATION

Date of Birth: October 17, 1968

Marital Status: Single

Business Address: Psychology Department
Virginia Polytechnic Institute
and State University
Blacksburg, VA 24061-0157

Business Phone: (540) 231-6581

Home Address: 770 Triangle St. #414
Blacksburg, VA 24060

EDUCATION

B.S. Kalamazoo College, 1991
Major field of study: Psychology

M.S. Virginia Polytechnic Institute and State University,
1996
Major field of study: Industrial/Organizational
Psychology

THESIS: EXAMINING THE IMPACT OF IMPRESSION MANAGEMENT
CONTEXT AND SELF-MONITORING ON THE LENIENCY AND
ACCURACY OF SELF-APPRAISALS

Major advisor: Dr. Neil M. A. Hauenstein

ACADEMIC EXPERIENCE

1996-1997 Graduate Research Assistant- Center for
Applied Behavior Systems, Department of
Psychology, Virginia Polytechnic Institute
and State University. Project coordinator for
numerous local industries, helped secure over
\$1.5 million dollars in grant money for the
fiscal year, mentored undergraduate
psychology students in data collection, data
entry, and conference presentation
preparation, developed and implemented
behavior-based observational checklists in
local industries.

1995-1996 Graduate Assistant- Department of Psychology, Virginia Polytechnic Institute and State University. Responsible for test development and grading for two undergraduate classes.

1994-1995 Graduate Lab Instructor- Department of Psychology, Virginia Polytechnic Institute and State University. Responsible for teaching two Introductory Psychology undergraduate labs. Conducted lectures and prepared and graded exams.

APPLIED EXPERIENCE

1996-1997 Safety Consultant- Safety Performance Solutions, Blacksburg, Virginia. Conducted behavior-based safety workshops in industries nationwide, worked in conjunction with organizations' safety committees, presented behavior-based safety research at safety conferences nationwide.

1993 Standardized Test Evaluator- The Psychological Corporation, San Antonio, Texas. Scored and graded standardized high school achievement tests and essays.

1991-1992 Lead Child Care Counselor- Lakeside Boys and Girls Residence, Kalamazoo, MI. Developed and implemented behavioral program designed to facilitate appropriate social behavior for delinquent and emotionally impaired adolescents. Served as a liaison between counselors and management.

1990 Head Tennis Pro- John Newcombe's Tennis Resort, New Braunfelds, Texas. Coordinated adult and children's tennis clinics in Mexico and the United States.

1989 Research Assistant- University of Texas Health Science Center, San Antonio, TX. Examined the effects of reward/punishment manipulations on the motor task efficiency of Attention Deficit Hyperactivity Disorder (ADHD) children. Neurochemical differences

between ADHD and control groups were also investigated.

RESEARCH PRESENTATIONS

1997 SOCIETY OF INDUSTRIAL ORGANIZATIONAL PSYCHOLOGY CONFERENCE, St. Louis, MO.

Joshua H. Williams, Charles B. Pettinger, Jr., Thomas E. Boyce, and Jason N. Fortney.

Participative versus non-participative approaches to safety training: Comparative impact on satisfaction and information retention.

1997 SOCIETY OF INDUSTRIAL ORGANIZATIONAL PSYCHOLOGY CONFERENCE, St. Louis, MO.

Charles B. Pettinger, Jr., Jason P. DePasquale, Thomas E. Boyce, Joshua H. Williams, and E. Scott Geller.

Critical success factors for increasing safe work practices: A systematic evaluation of real-world applications.

1997 SOCIETY OF INDUSTRIAL ORGANIZATIONAL PSYCHOLOGY CONFERENCE, St. Louis, MO.

Jason DePasquale, Kent E. Glindemann, Jason N. Fortney, Joshua H. Williams, and Amy B. Gershenoff.

Relative application and perceived value of various approaches to improve occupational safety: A nationwide survey of safety professionals.

1997 SOUTHEASTERN PSYCHOLOGICAL ASSOCIATION, Atlanta, GA.

Joshua H. Williams, Charles B. Pettinger, Jr., Thomas E. Boyce, and Jason N. Fortney.

Participative versus non-participative safety training: Information retention versus satisfaction.

1997 SOUTHEASTERN PSYCHOLOGICAL ASSOCIATION,
Atlanta, GA.

Charles B. Pettinger, Jr., Joshua H.
Williams, Thomas E. Boyce, and Deborah K.
Ford.

*DO IT for safety: A process of continuous
improvement.*

**1997 AMERICAN INDUSTRIAL HYGIENE CONFERENCE
AND EXPOSITION,** Dallas, TX.

Joshua H. Williams

*Participative versus non-participative safety
training approaches: Information retention,
perceived involvement, and satisfaction with
training.*

**1997 ASSOCIATION FOR BEHAVIOR ANALYSIS
CONFERENCE,** Chicago, IL.

Candice R. Chevaillier, Joshua H. Williams,
Paul G. Michael, Charles B. Pettinger, Jr.,
and Thomas E. Boyce.

*Involve them and they'll understand: a
systematic test of this training slogan.*

**1996 SOCIETY OF INDUSTRIAL ORGANIZATIONAL P
PSYCHOLOGY CONFERENCE,** San Diego, CA.

Steve Morgan, Josh Williams, and Dr. Neil
Hauenstein.

*Examining the effects of social comparison
information and expectation of validation on
the accuracy and leniency of self-rated
performance.*

1996 BEHAVIOR NOW CONFERENCE, Houston, TX.

Steve Roberts, Kent Glindermann, and Josh
Williams.

*The Use of a Safety Culture Survey
to Supplement a Behavior Based Safety
Process.*

PROFESSIONAL AFFILIATIONS

American Psychological Association
Society for Industrial/Organizational Psychology

RELEVANT GRADUATE COURSES

Research Methods
Industrial Psychology I, II
Organizational Psychology I, II
Theories in Personality
Quantitative Topics I, II
Psychometrics
Statistics for Social Sciences I, II
Social Psychology
Multiple Regression
GTA Training Workshop
Cognitive Psychology
Reward and Compensation Systems

COMPUTER SKILLS

Experience using SPSS, SAS, MS-DOS, Windows, Word, Excel, Powerpoint.

A handwritten signature in black ink, appearing to be 'J. M. ...', written in a cursive style.