Organizational Behavior Research
Department of Management
Department of Psychology

Texas A&M University
SELF AND SUPERIOR ASSESSMENT

Cynthia D. Fisher
and
Gail Russ

June, 1986
TR-UNR-4

Department of Management
Texas A&M University

Prepared for:
Office of Naval Research
300 N. Quincy Street
Arlington, Virginia 22217

This report was prepared for the Manpower R&D Program of the Office of Naval Research under contract N00014-85-k-0283. Reproduction in whole or in part is permitted for any purpose of the United States Government.
This theoretical paper begins by documenting the severity and pervasiveness of disagreements between superior and self assessments of work performance. To understand this disagreement, theoretical and applied work on self assessment is needed, similar to the work on superior cognitive processes during rating which has been occurring for several years. This paper draws on several theories to explain the self evaluation process and contrast it to the process used by superiors. A model of self evaluation, future research needs, and some suggestions for improving agreement are presented.
Performance appraisal research has always had the goal of developing appraisal instruments and procedures which would minimize distributional and intercorrelational error, and maximize rating accuracy. Inter-rater agreement was considered to be one of several important indicators of rating instrument quality. When various raters, such as self and superiors, could not agree, researchers generally blamed the format of the instrument and so the search for an ideal format continued. More recently, researchers have begun to focus on the cognitive and motivational processes that underlie performance rating, and have realized that some disagreement between raters at different levels is to be expected (Borman, 1974; Landy & Farr, 1980).

In this paper, we will explore this idea further. Specifically, we will begin by documenting the existence, magnitude, and nature of disagreement between superior and self ratings of performance. Then we will explore several perspectives on self ratings and attempt to construct a model of the self rating process. Self rating processes will be compared to superior rating processes to achieve a greater understanding of why the two sources disagree and what might be done to enhance agreement.

Self vs Superior Ratings—Evidence of Disagreement

A number of studies have found mean differences when comparing ratings of subordinate performance made by the subordinates themselves and by superiors. Subordinates usually, but not always (c.f. Heneman, 1974), rate themselves as better performers than do their superiors. For instance, Kirchner (1965) found that technical personnel rated themselves more favorably than did their superiors on seven out of seven performance dimensions, while Holzbach (1978) found that self ratings were significantly more lenient than superior or peer
ratings on four of seven dimensions. Klimoski and London (1974) asked a sample of nurses to rate themselves on 20 performance dimensions, and found that their mean rating was significantly higher than the mean rating obtained from their superiors. Prien and Liske (1962) reported similar results for a sample comprised of respondents from several organizations, job levels, and job titles. Thornton (1968) found that even executives rated themselves significantly higher, on the average, than did their bosses. Shapiro and Dessler (1985) found that 65% of lower level supervisors rated their performance as outstanding or highly satisfactory, while their superiors rated only 23% of them in those categories. Finally, Shore and Thornton (1986) reported that regardless of gender of ratee and rater, subordinates rated themselves higher on all dimensions than did their superiors.

Additional evidence for superior-subordinate disagreement on subordinate performance levels comes from correlational studies. Superiors and subordinates disagree not just on mean levels of performance, but also disagree in an ordinal sense, on how subordinates perform relative to each other. For example, Baird (1977), Brief, Aldag, and Van Sell (1977), Heneman (1974), Holzbach (1978), Klimoski and London (1974), Lawler (1967), Prien and Liske (1962), and Thornton (1968) all found mean or median correlations of superior-subordinate ratings (across subordinates on a variety of performance dimensions) lower than .27. A few studies have found slightly greater agreement--Pym and Auld (1965) reported inter-rater agreements on a single overall performance question ranging from .42 to .69 in samples of several types of employees. Parker, Taylor, Barrett, and Martens (1959) reported superior-subordinate inter-rater reliabilities of .13 to .53, depending on what aspect of performance was being rated. The greatest disagreement occurred when "leadership potential" was rated.
The Parker et al. study brings up a third aspect of superior-subordinate disagreement—that the amount of disagreement can vary with the performance dimension being assessed. Heneman (1974) found significant, though modest, correlations between self and superior ratings of planning, supervising, negotiating, representing, and overall performance, while non-significant correlations between raters appeared for investigating, coordinating, evaluating, and staffing. Thornton (1968) reported that superiors and subordinates disagreed the most when rating decisiveness, analysis, and development of subordinates. Finally, Williams & Seiler (1973) found that superiors and subordinates agreed more when rating performance than when rating effort, while Lawler (1967) reported just the opposite.

A fourth point about the nature of superior-subordinate disagreement on subordinate performance involves the extent to which the two types of raters commit halo error. "Halo error" is considered to exist when intercorrelations between ratings on different dimensions of performance are too high. Several researchers have discovered that superiors display slightly more halo error than subordinates (Baird, 1977; Heneman, 1974; Kirchner, 1965; Parker et al., 1959; Thornton, 1968; Williams & Seiler, 1973). Subordinates apparently can distinguish between their performance on different dimensions more clearly than can their superiors.

In conclusion, it seems that superiors and subordinates typically experience substantial disagreement when both rate the subordinate's job performance. Further, this disagreement may be manifested in several ways: (1) mean ratings by subordinates which are different (usually higher) than superior ratings, (2) very low intercorrelations between ratings made by superiors and subordinates, (3) more severe disagreement (both in terms of means and correlations) on some performance dimensions than others, and (4) greater halo error in superior ratings than in subordinate self ratings.
However, there do seem to be conditions under which superior and self assessments converge to a greater or lesser than usual degree. Perhaps by examining these conditions, some insight can be gained into the causes of the disagreement which so often occurs. Brief, Aldag, and Van Sell (1977) proposed ten demographic and attitudinal moderators of the level of agreement between self and supervisory ratings. They found that only one, tenure, was a significant moderator. Low tenure employees (less than 2 years) agreed with their bosses' assessments more than higher tenure employees. Herold and Parsons (1980) also reported a negative relationship between agreement and subordinate tenure. In addition, these researchers found that agreement was better when superiors and subordinates had a "high quality relationship" as indexed by subordinate satisfaction with the superior and perceived leader consideration, participativeness, and influence in the organization. Baird (1977) found that the greatest disagreement occurred among high self esteem subordinates whose superiors rated them low on the performance measure. Similarly, Thornton (1968) found that individuals rated as least promotable by a panel of superiors had the greatest disagreement with their immediate superior on 27 trait ratings, with subordinates consistently over-rating themselves. Shapiro and Dessler (1985) reported that disagreement on performance level was greatest between supervisors with a high school degree or less and their superiors. Disagreement was less severe in samples of college educated and graduate degreed supervisors and their superiors.

The self rating of ability literature may also contain some clues on agreement. When comparing self ratings of ability to either superior ratings or objective measures, agreement seems to be higher when subordinates are experienced in self evaluation, realize that their ratings may be compared
with objective measures, compare themselves to others rather than to an absolute standard, and are of high intelligence, internal locus of control, and high achievement status (Mabe & West, 1982).

Somehow, this recitation of conditions under which agreement may be enhanced is not very satisfying. There is no coherent framework uniting the various moderators of agreement, and even when favorable values of several moderators are present, interrater reliability is still disappointingly low. What is needed is a better understanding of the processes underlying disagreement. This can only be obtained by going beyond the disagreement itself to consider the origin of the two distinct types of rating. Heneman (1980, p. 298) states that, "Why discrepancies exist among self, peer, and supervisory ratings is not really known, and is very unlikely to be known, until a theoretical foundation is advanced for self assessment in performance appraisal." Recent research on schematic and attributional effects in supervisory rating have expanded our understanding of that process (Feldman, 1981; Ilgen & Feldman, 1983; Mitchell & Wood, 1980; Nathan & Alexander, 1985.) It is time to develop similar insights into the process of self assessment. An understanding of the two respective processes will allow for comparison between them and for the prediction of when and how differing perceptions will exist.

There are several well known theories which are concerned with self evaluation in general. They attempt to explain how self-concepts are formed, and how we come to know what kind of people we are. These will be presented in turn, and the particular self-evaluation process they describe contrasted with superior evaluation processes to see why they might produce divergent judgments. Eventually, insights from the various theories will be distilled into a model of the self-evaluation process as it differs from the evaluation process used by superiors.
Symbolic Interactionism

One view of self assessment is supplied by the symbolic interactionists (Cooley, 1902; Mead, 1934), who state that self perceptions are generated largely from individuals' experiences of how others view them. The term "looking glass self" has been used to convey the idea that the self concept simply reflects others' assessments. If this is correct, one would not expect to see much disagreement between self and superior ratings of performance, assuming that adequate feedback is given and the superior is a "significant other." However, disagreements clearly do exist. Ilgen, Peterson, Martin, and Boeschen (1981) found that disagreements remained even immediately after an appraisal interview in which a great deal of feedback was given.

Research on the symbolic interactionist perspective has shown that others' assessments can influence self perception, but are far from the only determinant of self concept. Kinch (1968) found that assessments by others had their greatest impact when given frequently. Sherwood (1966) found that individuals modified their self assessments to match those of others when the others were in agreement with each other. When there was disagreement among the others, focal persons tended to ignore the other evaluations and see themselves positively (Sherwood, 1966; Kinch, 1968). Stone and Stone (1985) offered feedback from two sources which did not agree. One source told the performer that his/her work was very good, the other said that the work was merely acceptable. Performers found the positive source to be more credible and utilized the positive rather than the neutral feedback in making a self assessment. Recent research on social influence has confirmed that the frequency, number, and consistency of sources affect their success in changing the perceptions of a target (Latane, 1981).
However, after a thorough review of the literature, Shrauger and Schoeneman (1979, p. 549) concluded that "...people's self-perceptions agree substantially with the way they perceive themselves as being viewed by others. However, there is no consistent agreement between people's self-perceptions and how they are actually viewed by others. There is no clear indication that self-evaluations are influenced by the feedback received from others in naturally occurring situations." Recently, Schafer and Keith (1985) have criticized this review and earlier work for not picking others that are truly significant as sources of feedback, and for attacking an overly simplified version of the symbolic interactionist model. They state that actual assessments by others should influence perceptions of these assessments, and that these perceptions in turn influence the self-concept. Thus, their model does not require a direct relationship between others' actual assessments and self-concept. In a path-analytic study of spouses' perceptions of each other and themselves, Schafer and Keith supported their view that: 1) others' assessments affect perception of others' assessments, 2) perception of others' assessments influence self concept, 3) there is virtually no direct influence of others' perceptions on the self concept.

Thus, it seems likely that superiors' assessments of performance will have relatively little relationship to subordinates' private views of their own performance level. The many low correlations reported in the introduction lend support to this assertion. According to Schafer and Keith (1985), what might influence self-evaluations are perceptions of the superior's evaluation. Unfortunately, the evaluations of others may not be accurately perceived. Smircich and Chesser (1981) found a correlation of .04 between superior ratings and subordinate estimates of superior ratings. Parker et al. (1959) had some what more promising results. They asked subordinates to rate their
performance as they saw it, then rate it again as they believed their superiors saw it. The usual leniency effect occurred, with self-ratings being quite a bit higher than superior ratings. Perceptions of the superiors' ratings were in between, such that they might have been influenced by actual superior assessments and subsequently influenced self assessments. These results are consistent with the revised symbolic interactionist perspective.

Table 1 summarizes the symbolic interactionist perspective and its implications and conclusions for superior subordinate disagreement.

Social Comparison Theory

A second view on the origin of self perceptions is that they are the result of an active comparison process between the self and others, as described in Festinger’s (1954) social comparison theory. Individuals desire to form accurate and stable self assessments of their opinions and abilities. In the absence of unambiguous objective standards, people must compare themselves to others in order to make these assessments. According to Festinger, similar others are most often chosen for comparison purposes. In the case of opinions or attitudes, similar others are likely to agree and verify that one’s views are correct, and it is this consensus that is desired in making opinion comparisons.

In making ability comparisons, a somewhat different set of motives seems to be operative. First, several studies have found a tendency for individuals to compare with others who are extreme or outstanding on the trait in question, rather than to compare to others nearer their own level (Gruder, 1977). This has been explained by Thornton and Arrowood (1966) as an attempt to make an accurate self-evaluation by comparing to a clear, positive instance
Table 1

Symbolic Interactionism

<table>
<thead>
<tr>
<th>Theory</th>
<th>Implications</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self perceptions are based on the communicated perceptions and evaluations of significant others. (not supported)</td>
<td>Subordinates should largely agree with their superiors, if adequate feedback is given. (not supported)</td>
<td>Feedback from the superior is not the only determinant of self-evaluation.</td>
</tr>
<tr>
<td>as revised:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self perceptions are based on how individuals believe they are viewed by others. Individuals do not always accurately perceive the way they are viewed by others. (some support)</td>
<td>Self-evaluations should be more similar to expected superiors' evaluations than to actual superiors' evaluations. (mixed support)</td>
<td></td>
</tr>
</tbody>
</table>
of the construct (trait, ability) in question, rather than to the quasi-instance of a more similar peer. Thus, desire for accurate self-evaluation may result in the choice of a superior other as a comparison.

Second, Festinger has suggested that in the case of abilities (but not opinions), individuals wish to find that they possess a good or high level of desirable abilities. However, if the most informative comparison is to the top performer, then this second motive will be frustrated, in that the individual will almost always be lower in ability than the comparison other. There is evidence that both motives, accuracy and self-enhancement, do operate and that each may take precedence at different times (Gruder, 1977).

A substantial body of literature suggests that individuals are motivated to form self-assessments that are accurate. Darley and Goethals (1980, p.11) point out that "... it is useful for people to know their own abilities... because it enables them to predict the success or failure of their efforts. [This] keeps people from attempting tasks that are too difficult, which keeps them from the normally negative consequences of failure." Trope, in a series of studies (1975; 1979; 1980; Trope & Brickman, 1975), has shown that when subjects are allowed to select their own test items, item diagnosticity is the strongest determinant of choice. This preference is enhanced when individuals are relatively more uncertain of their ability level (Trope, 1982, Trope & Ben-Yair, 1982). Finally, the preference for diagnostic tasks is especially strong in individuals high in resultant achievement motivation. In none of Trope's published research was there any tendency for subjects, even those induced to believe that they had low ability, to avoid diagnostic tasks. This is in startling contrast to numerous other studies in which active avoidance of diagnostic tasks has occurred, apparently in the service of self-enhancement rather than accuracy of assessment motives.
For instance, several studies have found that individuals avoid diagnostic tasks if they expect to do poorly on them (Conolley, Gerard, & Kline, 1978; Sachs, 1982; Zuckerman, Brown, Fischler, Fox, Lathin, & Minasian, 1979). Others who expect to do well in subsequent testing (as a result of credible positive feedback on earlier trials) show the usual preference for diagnostic items. In a study notable for its realism, Meyer and Starke (1982) offered students in a counseling and guidance course the opportunity to score and compare to normative data one of the two tests which they had already taken. One was an intelligence test, the other a preference inventory without correct answers. Several days earlier, students had also completed a "self-concept of ability" measure. Individuals in the top quartile on self rated ability overwhelmingly chose to score the intelligence test (72%), while those low in self rated ability showed an equally strong preference to score the non-performance test (67%). Thus, in a real-life setting, individuals chose to avoid making a comparison which they expected to yield negative feedback. Interestingly, there was no relationship between self-rated ability and objectively measured intelligence.

Returning to the literature on social rather than objective comparison, there is again some evidence that social comparisons can be intentionally skewed in the service of self enhancement. Hakmiller (1966) found that under a high threat to self esteem, subjects chose to compare themselves (favorably) to others who were markedly inferior on the relevant dimension. Friend and Gilbert (1973) found similar results, and also reported that individuals chronically high in fear of negative feedback (low in self-esteem) were especially likely to make defensive comparisons to worse-off others under high situational threat to esteem. Jones and Regan (1974) have helped to clarify when accuracy motivation will prevail, and when self-enhancement motives will
dominate the comparison process. They showed that accuracy is of concern when one expects to make a decision regarding future participation in an ability relevant activity. Accurate self assessment under these conditions may help one avoid embarrassment and maximize success. When future activity is not an issue, as in the Hakmiller (1966) and Friend and Gilbert (1973) studies, self-enhancement motives dominate.

Undoubtedly, some component of one's self evaluation comes from comparison with others. Employees are likely to choose from among their peers specific others with which to compare themselves (Adams, 1965). However, choice of others may be systematically biased so as to yield a positive comparison rather than an accurate one, particularly if the comparer feels threatened.

The superior as a rater will not succumb to this bias. Superiors often have available a large amount of relevant information against which to compare each subordinate, if they have observed numerous subordinates, both past and present, in the target job. On the other hand, superiors may not use the information potentially available to them. Instead, they may display the "false-consensus bias" (Hansen & Donoghue, 1977; Ross, Greene, & House, 1977) and use themselves for comparison, assessing each subordinate against how well the superiors think that they could have done the job. (This bias will be discussed more later in the paper.) Clearly, if superior and subordinate are using different comparison others, they are likely to reach different conclusions and disagree about the level of subordinate performance. See Table 2 for a summary of this section.

Consistency Theories

A third theoretical foundation for self assessment at work is provided by Korman (1970), drawing on older balance and dissonance theories. He
<table>
<thead>
<tr>
<th>Theory</th>
<th>Implications</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>People will compare themselves to objective standards when they are available. (supported)</td>
<td>Choice of different comparison others plus the subordinate's changing motivation to achieve an accurate or a favorable comparison is one reason for superior/subordinate disagreement.</td>
<td></td>
</tr>
<tr>
<td>In the absence of objective standards, comparisons are made to other people. (supported)</td>
<td>Individuals sometimes avoid obtaining diagnostic information or choose to compare themselves to worse off others in order to achieve a favorable comparison of abilities. (supported)</td>
<td></td>
</tr>
<tr>
<td>For opinion and attitude comparisons, similar others are chosen and consensual validation is desired. (supported)</td>
<td>In evaluating subordinates, superiors may compare to different others than subordinates compare themselves to.</td>
<td></td>
</tr>
<tr>
<td>For ability comparisons, individuals are both motivated to make accurate comparisons and to conclude that they possess a high level of desirable abilities. (supported)</td>
<td>Superiors may choose themselves or their own past performance (actual or imagined) as the comparison against which to evaluate subordinates' performance.</td>
<td></td>
</tr>
</tbody>
</table>
suggests that individuals develop a self-concept or self image concerning job performance. This self image is based on one's general level of self-esteem, one's previous success or failure on similar tasks, and the expectations and evaluations of others. Once formed, this self image influences the level of actual job performance, with individuals seeking to perform and receive feedback consistent with their self concepts. Thus, individuals with high job-specific self-esteem will prefer to succeed and will derive satisfaction from success, and individuals with low job-specific self-esteem will prefer to fail. It is this latter prediction which has caused the most problems for consistency theories.

Many studies have shown that people find consistent feedback to be more credible and memorable than inconsistent feedback (c.f. Shrauger, 1975). Swann and Read (1981a; 1981b) report six different studies supportive of the consistency view. Their subjects displayed clear preferences for feedback which was consistent with their self image, even when the trait in question was somewhat negative (unassertive). Subjects spent more time looking at consistent feedback, paid more money to obtain it, tried to elicit it in interpersonal situations, and recalled it better than inconsistent feedback. This would seem to provide strong support for the idea of a consistency motive. However, an alternative explanation is possible.

Swann and Read (1981a) also found that consistent feedback was perceived as being more informative and diagnostic. They suggest that information value may underlie all of the above manifestations of preference for consistent feedback. Trope (1979) would agree. He found that subjects who were led to believe that they were somewhere in the "good" range selected items which would further diagnose their exact standing in the range. Subjects in the "poor" range similarly selected items that were most diagnostic within their
range. Swann and Read's subjects who had assigned themselves to a particular range, such as more unassertive than assertive, were most interested in information which would allow a yet more precise self-evaluation within the range already known to apply. Thus, all these findings can be explained without the "strong" consistency hypothesis that individuals with low self-esteem prefer to fail (Dipboye, 1977). In fact, marked consistency-like effects have been observed in many setting in which self-esteem is not involved. For instance, Lord, Ross, and Lepper (1979) found that individuals who held strong positive or negative views about capital punishment became even more polarized after being exposed to ambiguous evidence on its effectiveness as a deterrent. Subjects accepted studies which supported their views uncritically but rejected disconfirming research as uncredible and fraught with methodological errors. The information processing basis for apparent consistency effects will be discussed in more detail later.

The implications of consistency theories for inter-rater agreement are clear. The subordinate will selectively seek and recall information consistent with his or her level of job-specific self-esteem, or current self-evaluation. The self-assessment should remain stable since only consistent information is considered diagnostic. The superior probably has little knowledge of the subordinate's self-esteem, and has no need to assess the subordinate in a manner consistent with that self-esteem even if it is known. The superior is likely to attend more widely to information about the subordinate and not exclude information which the subordinate would consider inconsistent. On the other hand, if the superior has formed an impression of the subordinate's level of performance, consistency-like biases may again come to bear as the superior ignores subsequent information which is inconsistent with the impression. This idea will be discussed more in the context of
schematic information processing. Thus, disagreement may persist as each rater rejects information discrepant with his or her initial and unchanging position.

Self-Enhancement Theories

Self-enhancement theories have evolved in response to some of the weaknesses of consistency theories, particularly the prediction that low self-esteem individuals prefer to fail. Jones (1973) and Dipboye (1977) reviewed evidence which showed that even persons of low self-esteem prefer success to failure. In fact, such persons are even more desirous of receiving positive or success feedback, because their needs for esteem are relatively less satisfied. The original studies which were taken as support of the consistency view that low self-esteem individuals seek failure have been reinterpreted to be consistent with a self-enhancement view (Jones, 1973). Actions which appeared to be failure-seeking behavior by low self-esteem individuals are now considered preemptive strikes in defense of what little self esteem they possess. Behaviors such as self-handicapping, not trying hard, or devaluing the task guarantee that any failures which may occur do not really reflect directly on the person and so do not endanger fragile self-esteem (Berglas & Jones, 1978; Jones & Berglas, 1978; Regan, Gosselink, Hubsch, & Ulsh, 1975).

Much of the psychology literature indicates that people desire to see themselves favorably as competent human beings (Bandura, 1982; White, 1959). At times, this may require holding an inflated image of one's self. Greenwald's (1980) review supports this conclusion. He presents evidence that people selectively recall and even rewrite their memories of past events so as
to maintain a positive self-image. Gollwitzer, Wicklund, and Hilton (1982), in their work on "symbolic self-completion," observed this process in action. They found that individuals lacking in unambiguous objective qualifications, or those induced to think about past failures in an area, were especially anxious to extoll their own successes in that area, and so to "complete" their self-images as competent people. Greenberg and Pyszczynski (1985) found that subjects given public failure feedback (and not allowed to derogate the test on which they had failed) responded with a large increase in self-regard, which the authors labeled "compensatory self-inflation. Further, Lewinsohn, Mischel, Chaplin, and Barton (1980) found that normal, well-adjusted people evaluated themselves significantly more positively than observers, whereas clinically depressed people evaluated themselves accurately. They concluded that, "To feel good about ourselves we may have to judge ourselves more kindly than we are judged" (p. 212). The attribution literature, which will be discussed shortly, also supports the idea that people strive to maintain positive self-images.

Shrauger (1975) thoroughly reviewed the evidence for consistency versus self-enhancement views, and found that self-enhancement best explained affective responses, in that positive evaluations produced greater satisfaction with the rating, especially for people with low self-esteem. Jones' (1973) review reached the same conclusion. On the other hand, Shrauger found that consistency theories were supported when cognitive reactions, rather than emotional ones, were considered. Feedback inconsistent with the self-concept was recalled less accurately, perceived as less credible, and accepted less as resulting from any enduring internal characteristics. These effects were found for individuals with low self-esteem who were given positive feedback as well as for high self-esteem people given negative
feedback. This phenomenon can be explained as an information processing problem rather than a motivationally based desire to achieve consistency even at the cost of failing.

The literature on self-presentation and impression management states what may be considered a corollary of self-enhancement theory: that individuals strive to make a good impression on others. Often this is achieved by presenting oneself positively, as a skilled and competent person. At other times, a display of modesty and self-depreciation may be more useful in producing a positive impression in the perceiver.

In the work setting, and even in laboratory research, it is methodologically difficult to disentangle true, private beliefs about the self from efforts to convey a particular impression to an audience, whether that audience is the superior, peers, or just the researcher. Further complicating the issue is the finding that presenting oneself positively to an audience (at the request of an experimenter) has a subsequent positive effect on private self-esteem (Jones, Rhodewalt, Berglas, & Skelton, 1981). Thus, impression management attempts aimed at the superior may also affect the subordinate’s own view of his or her ability.

Successful impression management by the subordinate would have the effect of reducing disagreement between ratings, because the superior would see the subordinate as the latter wished to be seen. However, it is unlikely that a subordinate could be completely successful in modifying a superior’s view, as the superior usually has other sources of reliable information. In fact, the subordinate may not try to engage in impression management on dimensions for which more objective information is available to the superior. Several studies have shown that individuals present themselves accurately to others who have objective knowledge of their weaknesses (Baumeister & Jones, 1978;
Schlenker, 1975; Ungar, 1980). On the other hand, when constrained to be accurate on known dimensions, individuals may be more motivated to convey a positive impression on other dimensions:

"...when a person realizes that someone else has an unfavorable impression of several of his or her personal characteristics (derived from an authoritative source), the person will not attempt to challenge that impression directly but will compensate for it by presenting him/herself all the more favorably in areas about which the other lacks information." (Baumeister & Jones, 1978, p. 616).

As noted in the introduction, superiors and subordinates disagree to different extents on different dimensions. It would be interesting to see whether dimensions with high agreement are those on which the superior lacks information sources other than the subordinate. If so, then one might conclude that the subordinate is managing the superior's impression effectively on those dimensions. Table 3 summarizes the points made in the sections on consistency and self-enhancement theories.

**Attribution Theory**

Attribution theory has a great deal to say about how people perceive themselves and others. There appear to be reliable biases in the way that people attribute causality for their own acts and those of others. Some of these biases clearly operate in the service of self-enhancement. For instance, the "actor-observer bias" occurs as observers consistently attribute causality to internal-to-actor factors. Whether an event is good or bad, it typically is seen by observers as being the actor's fault. Observers presumably have no general need to see actors as competent or successful. Actors, however, make quite different attributions, tending to claim causal credit for positive outcomes and deny blame (make situational attributions) for negative outcomes (Jones & Nisbitt, 1971). Similarly, when working in groups, individuals see themselves as disproportionately responsible for the
<table>
<thead>
<tr>
<th>Consistency Theory</th>
<th>Implications</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals desire to perform in a manner consistent with their positive or negative self-image. Those with a low self-image prefer failure to success. (little support)</td>
<td>Individuals selectively seek and recall feedback which is consistent with their existing self assessment. (supported)</td>
<td></td>
</tr>
<tr>
<td>Individuals find consistent feedback to be more credible and memorable than inconsistent feedback, perhaps because it is considered more informative and diagnostic. (supported)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Self-Enhancement Theory**

| Individuals desire to have a positive self-image. (supported) | Self-assessments of all types will tend to be lenient. (supported) | Subordinates will often over-estimate their own performance compared to objective or superiors' assessments. (supported) |
| People prefer success to failure. (supported) |  |  |
| People strive to make a good impression on others when possible. (supported) | Subordinates may attempt to engage in impression management with their superiors. | To the extent that impression management activities are successful, superior/subordinate disagreement may be enhanced. |
success of the group (Ross, 1981). There has been an ongoing debate as to whether the actor bias is due to information processing quirks or to motivational factors (see Miller & Ross 1975; Bradley, 1978). Several recent studies seem to indicate that actor bias is motivationally based -- claiming credit for success occurs for self-enhancement reasons, and rejecting blame for failure serves an ego-defensive function (Snyder, Stephan, & Rosenfield, 1978). Miller (1976) supported this view by finding that actors' tendencies toward asymmetrical attributions for success and failure are stronger when the task is portrayed as important and socially valued rather than unimportant. If attributions were simply an artifact of information processing, then task importance would not have this effect.

Sicoly and Ross (1977) had confederate observers allocate either more responsibility for success and less for failure, or less responsibility for success and more for failure to actors than the actors had attributed to themselves. If actors reached their own attributional conclusions via simple information processing, then they should see equal deviations on either side of their own estimates as equally inaccurate. In fact, actors felt that observers who gave more credit for success and less for failure were much more accurate than those who erred in the opposite direction, again lending support to a motivational explanation for actor bias.

However, stating that the bias is motivationally based is not the same as claiming that it is conscious or intentional. The fact that the bias is so pervasive may argue that it is unintentional, that individuals in western cultures learn very early to feel responsible for important successes and see plausible external causes for otherwise threatening failures. Thus, the motivationally-based bias may be an "honest mistake" of which the actor is unaware.
Clearly, attributional biases could account for a large part of the disagreement between self-and superior-assessments of performance. Even when a subordinate is trying hard to make an accurate self-assessment, attribution biases may affect the conclusion. The conscientious self-rater should first search memory for all relevant indicators or incidents of performance. Those for which one was not responsible will not be seen as relevant, and will not be weighted in the final judgement. Thus, a self-assessment will be based largely on instances of positive performance. Superiors, who tend to attribute all outcomes, whether positive or negative, to the actor, should naturally produce a somewhat lower performance judgement. This tendency is further exacerbated by superiors' predisposition to make stronger internal-to-subordinate attributions for a behavior when the behavior leads to a negative outcome (Mitchell & Kalb 1981; Mitchell & Wood, 1980). Ferris (1984) verified that when superiors and subordinates agree about the causes of performance, feedback is seen as more accurate and fair. Subordinates had the expected preference for internal attributions following good performance and external attributions following poor performance.

One factor which may bring the superior's judgment into closer accord with the subordinate's view is whether or not the superior has previously performed the subordinate's job. Mitchell and Kalb (1982) conducted a lab study investigating the impact of previous experience on supervisor's ratings and attributions regarding a poor performing subordinate. They predicted, and found, that superiors who have themselves been "actors" before becoming "observers" tended to make more actor-like (situational) attributions for the poor performance of their subordinates. An interview study of Army officers also confirmed the tendency to make external attributions for poor subordinate performance when officers had previous experience in their subordinates' jobs (Mitchell and Kalb, 1982).
Additional research is needed on whether a similar phenomenon would occur in evaluating good performers. If having been an actor primes actor-like attributions, then good performers should be rated more highly and given more credit for causing their own success by experienced than unexperienced superiors. Verification of these results in a field setting is also needed.

An attribution issue not addressed by Mitchell and Kalb (1982) but which could be quite important in producing superior subordinate disagreement is the self-assessed performance level of the superior when he/she held the subordinate’s job. Or, if the job was not actually held, the superior’s beliefs about how well he or she could have performed the job. Social psychologists have documented a tendency to view one’s own attitudes and behavior as normative, and to evaluate and make attributions about others based on their similarity to oneself (Hansen & Donoghue, 1977; Ross, Greene, & House, 1977). This is called the "false consensus" effect. Specifically, when others behave similarly to the evaluator, their behavior is seen as being caused by the same reasonable external conditions that caused the evaluator’s behavior. When others behave quite differently, they are seen as being odd, or having extreme personal dispositions which cause their behavior to differ from that of a reasonable person.

If this tendency occurs in performance rating, the superior might make dispositional attributions and give more extreme ratings to subordinates who perform differently (both better or worse) than he or she did. However, the research on false consensus which led to this prediction has been limited to opinion or choice situations in which ability is not an issue. Given the desire to achieve superiority and competence rather than consensual validation when making ability comparisons, the effect of previous performance on evaluations of subordinates may be different. Specifically, it may be
asymmetrical, with differences between one's own performance and that of less successful subordinates being magnified and attributed to internal causes, while differences between oneself and better performing subordinates are minimized and attributed to a facilitating environment. In other words, "the subordinate who is worse than me is very poor, mostly because he/she is stupid or untalented. The one who is better than me isn't really that much better, and just got lucky."

Any of these mechanisms could create additional superior/subordinate disagreement, as the superior's judgment is being colored by factors of which the subordinate is unaware, and which he or she would consider irrelevant even if known. To reiterate, the superior's previous performance level on the subordinate's job, or imagined performance level when there is no previous experience, may bias both rating level and attributions in a way that magnifies superior/subordinate disagreement (see Table 4).

Schema Theory

A quite different and more general approach to understanding self-concepts and self-assessments that has gained much attention recently concerns the role of schemas in information processing. A schema is a cognitive structure that organizes information on attributes and interrelationships of attributes of a particular stimulus or concept (Fiske & Taylor, 1984). Schemas guide the processes of perceiving new information, storing and retrieving data, and making inferences. Essentially, schemas serve as frameworks which help organize and categorize information, and as a result of the imposed categorization, determine what is stored in memory and what is later available for making inferences (Markus & Sentis, 1982).
### Table 4

**Attribution Theory**

<table>
<thead>
<tr>
<th>Theory</th>
<th>Implications</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observers tend to make internal attributions, especially for poor performance. (supported)</td>
<td>Each evaluator will weigh only those performance incidents considered to be internally caused by the actor.</td>
<td>Commonly occurring attribution biases tend to increase superior/subordinate disagreement.</td>
</tr>
<tr>
<td>Actors tend to make external attributions for failure and internal attributions for success. (supported)</td>
<td>Superiors and subordinates may often disagree about causality, and hence about which performance incidents are relevant.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Previous experience by the superior in the target job may lead to more actor-like attributions. (supported)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>False consensus bias may result in inappropriate and ego defensive attributions by the superior.</td>
<td></td>
</tr>
</tbody>
</table>
Self-schemas form the basis of the concept of "self" by supplying the individual with well developed generalizations about the kind of person he or she is. Not everyone has the same set of schema structures. Individuals develop self schemas for dimensions in which they consider themselves extreme (e.g., very aggressive or very trustworthy) and are aschematic for dimensions which they consider that they possess in only a moderate or slight degree (Markus, 1977). Further, people tend to rate the traits on which they are schematic as highly important, and to see aschematic traits as unimportant (Markus & Smith, 1981).

Markus (1977) studied the effects of self schemas on processing information about the self. Her subjects were individuals who considered themselves independent or dependent (schematics) or who did not have a strong feeling that either extreme was self-descriptive (aschematics). Markus found that the "independent" schematics could make judgments about whether or not independence-related adjectives characterized themselves very quickly, whereas they took considerably longer to decide about dependence-related adjectives. "Dependent" schematics showed the opposite effect: their decisions were made much more quickly for dependence-related adjectives than for independence-related adjectives. The aschematics showed no difference in their processing time for the two types of adjectives. Markus interpreted her finding as demonstrating that self-schemas allow for faster processing of schema-related information. She also found evidence that self-schemas allow the person to recall more schema-consistent behavior. Independent schematics were able to remember more examples of their independent behavior than were the other two groups and dependent schematics could recall more examples of dependent behavior. Markus also found that independent schematics were resistant to information that suggested they were not independent, as were
dependent schematics about information that they were not really dependent. As schematics, on the other hand, seemed to be more accepting of such information.

Several researchers have found that individuals have better memory of events when those events can be compared to self-schemas (Bower & Gilligan, 1979; Kuiper & Rogers, 1979; Rogers, Kuiper, & Kirker, 1977). For instance, Bower and Gilligan found that subjects were able to remember trait adjectives better when they judged the adjectives in relation to themselves, rather than judged them for meaning or sound. They also demonstrated that lists of traits were recalled better when the subjects considered them in reference to themselves or to their mothers rather than to an unfamiliar person. They conclude that "good memory depends on relating the inputs to a well-differentiated memory structure" (p. 420). As Markus (1977, P. 63) points out, a "substantial amount....some might even argue a majority of the information we process is information about the self, so self schemas should be the most complex and well-differentiated memory structures we possess."

Certainly in the area of work behavior and performance, individuals have a great deal of information about themselves. Work is also a major component of most people's lives and therefore very important. These two factors--amount of information and importance--suggest that individuals will probably develop quite detailed and complex schemas for storing information and making self assessments of job performance. Schemas should exist for dimensions of work performance which incumbents believe are important and perhaps on which they feel they are extreme. Such dimensional schemas might include four components: a definition of the dimension, remembered examples of one's behavior relevant to the dimension, remembered feedback from others about one's standing on the dimension, and a self assessment of one's standing on the dimension.
Purposes of Schemas

A schema serves several purposes in assessment. First, it provides a framework for remembering information about performance. The schema provides labeled, predefined cubbyholes into which new information may be slotted. More complex schemas have more cubbyholes and can hold more information than less complex schemas. Self schemas tend to be very complex, as each individual has spent his or her life in observing, studying, and making sense of the self. Thus, it seems likely that an individual would have much more performance information stored in his or her well developed self-schema than would a superior in his or her schemas about subordinates.

Not much is known about the schema structures used by superiors in rating subordinates. However, there is some evidence that they use both global performance level schemas (Feldman, 1981) and more specific dimensional schemas (Borman, 1978). Upon first impression, a subordinate may be assigned to a superordinate category such as "good performer", "average performer", or "poor performer" (Cantor & Mischel, 1979). This initial categorization has some lasting effects. It accounts for halo error and it induces "false positive errors" - recalling that a subordinate performed a category-consistent behavior when in fact he or she did not (Major & Foti, 1985; Nathan & Lord, 1983; Ostrom, Lingle, Pryor, & Geva, 1980). DeNisi, Cafferty, and Meglino (1984) suggest that initial assignment to a category influences the perception and labeling of subsequent behavior, such that a "good performer" observed taking a long coffee break is considered to be reorganizing his or her thoughts before plunging into the next project, while a "poor performer" engaging in the same objective behavior is considered to be lazy.
There is also solid evidence that superiors are capable of using dimensional schemas in which individual behaviors are stored and averaged to produce a rating on the dimension. This rating is minimally affected by the general impression categorization, at least under laboratory conditions in which competing information processing demands are minimal (Major & Foti, 1985; Nathan & Lord, 1983). Figure 1 represents diagramatically a memory structure including both global and dimensional schemas which could be used by a superior.

A second purpose served by schemas is to direct attention to the specific dimensions on which one is schematic. Information or behaviors relevant to a dimension on which one is aschematic may not be perceived. If perceived, they may not be stored in memory, since there will not be a preexisting cubbyhole prepared to receive this type of information.

If superiors and subordinates are schematic on different dimensions, they might notice, remember, and value entirely different sorts of behavior in evaluating subordinate performance. There is some evidence that superiors and subordinates do hold different schemas for the subordinate's job. For instance, Borman (1974) has found that superiors and subordinates disagree on which dimensions should be included in behaviorally anchored rating scales for the latter's job. Both Zammuto, London, and Rowland (1982) and Schmitt, Noe, and Gottschalk (1986) have demonstrated that superiors and selves differ in the weights given to various performance dimensions in arriving at an assessment of overall performance. Holzbach (1978) factor analyzed an 18 X 18 correlation matrix made up of ratings on six performance dimensions from three sources: self, peer, and superior. The first three factors to emerge, accounting for 65% of the total variance, were very clear rater factors rather than performance dimension factors. Klimoski and London (1974) used a similar
Figure 1

Superior’s Schema Structure

Superordinate Categories

<table>
<thead>
<tr>
<th>Good Performer</th>
<th>Average Performer</th>
<th>Poor Performer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary</td>
<td>Fred</td>
<td>Susan</td>
</tr>
<tr>
<td>Bob</td>
<td>Terry</td>
<td>Lee</td>
</tr>
</tbody>
</table>

Dimensional Categories

<table>
<thead>
<tr>
<th>Productive-Unproductive</th>
<th>Good-Poor Teacher</th>
<th>High-Low Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary</td>
<td>Susan</td>
<td>Terry</td>
</tr>
<tr>
<td>Bob</td>
<td>Lee</td>
<td>Bob</td>
</tr>
<tr>
<td>Fred</td>
<td>Terry</td>
<td>Fred</td>
</tr>
<tr>
<td></td>
<td>Mary</td>
<td>Susan</td>
</tr>
<tr>
<td></td>
<td>Lee</td>
<td>Fred</td>
</tr>
</tbody>
</table>

method and also discovered distinct rater bias factors for each source of ratings. This research also supports the idea that superiors and subordinates have systematically different views of subordinate performance. Finally, Bernardin and Villanova (1984) took a more direct approach by simply asking raters and ratees to comment on the extent to which a variety of factors contributed to inaccuracy in superior ratings. Subordinates believed to a much greater extent than superiors that the latter attached too much weight to unimportant dimensions, did not consider all aspects of the job, and did not consider factors beyond the subordinate's control which may influence performance. Thus, there does seem to be evidence that superiors and subordinates are schematic on different dimensions of job performance.

A second aspect of the attention-directing function of schemas is the well documented tendency to evaluate others with the same dimensions on which one is strongly self-schematic. In one early study, subjects picked the ten dimensions out of fifty seven (such as fat-thin, sociable-unsociable) which they felt were the most and least relevant in evaluating themselves. Sometime later, they were asked to write free descriptions of other people they knew. The self-relevant dimensions were applied much more frequently in describing others than the non-self-relevant dimensions (Shrauger & Patterson, 1974). Lewicki (1983) has researched what he calls the "self-image bias in person perception," and confirmed that the dimensions on which one rates oneself highly also tend to be central in one's descriptions of others. Two recent reviews summarize similar research indicating that self-schemas do influence the dimensions people attend to in perceiving others (Markus & Sentis, 1982; Markus & Smith, 1981). Figure 2 demonstrates both the greater complexity of self-schemas and the tendency to evaluate others on the dimensions used most in self-evaluation.
Figure 2

Subordinate's Schema Structure

<table>
<thead>
<tr>
<th>Good Performer</th>
<th>Good Performer</th>
<th>Average Performer</th>
<th>Poor Performer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligent</td>
<td>Mary</td>
<td>Susan</td>
<td>Lee</td>
</tr>
<tr>
<td>Grant getter</td>
<td>Bob</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insightful</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good writer</td>
<td></td>
<td>Intelligent-Stupid</td>
<td>Modest-Prima Donna</td>
</tr>
<tr>
<td>Tough on students</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrious</td>
<td>Mary</td>
<td>Lee</td>
<td>Mary</td>
</tr>
<tr>
<td>Well organized</td>
<td>Bob</td>
<td>Fred</td>
<td>Bob</td>
</tr>
<tr>
<td>Good teacher</td>
<td>Susan</td>
<td></td>
<td>Susan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Appropriately Modest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respected</td>
</tr>
<tr>
<td>Not a leader</td>
</tr>
<tr>
<td>Insufficiently appreciated</td>
</tr>
<tr>
<td>Reliable</td>
</tr>
<tr>
<td>Mediocre statistician</td>
</tr>
<tr>
<td>Team player</td>
</tr>
</tbody>
</table>
It seems that superiors are likely to judge subordinates on the dimensions which they use most in judging themselves. One superior may be strongly self-schematic on being a classy dresser and being scrupulously honest, while another assesses him or herself primarily on being well organized and punctual. Each would tend to form an impression of subordinates based on these idiosyncratic dimensions. The dimensions on which a superior is schematic may or may not have true relevance to the subordinate's job, and may or may not co-occur in the subordinate's self-schema. To the extent that different dimensions are relevant to the two parties, their assessments of the subordinate's work behavior are likely to disagree.

One factor which may increase the similarity of superior and subordinate schemas is prior experience by the superior in the subordinate's job. While an incumbent, the superior-to-be would develop a more detailed schema for various aspects of the job. This complex structure might later be employed to understand and remember subordinate behavior on the job. As mentioned earlier, Kalb and Mitchell (1982) found that previous incumbency led supervisors to make more actor-like attributions for subordinate poor performance. Using a schema rationale, we also would predict that such superiors would be able to recall more subordinate behavior than would superiors without relevant job experience. This greater schema complexity and information availability could lead to less halo error, more accurate ratings, and possibly to ratings which agree more closely with subordinates' self-assessments.

A third function of schemas is to censor incoming information which would tend to contradict an established schema. Basically, schemas resist change. If they changed easily in response to all new or discrepant information, then they would not be able to serve their function of helping to simplify and
impose some stability on a complex information environment (Crocker, Fiske, & Taylor, 1984). Of course, schemas can and do change somewhat in response to enough discrepant information.

It appears that the type and amount of schema change depends partly on the complexity of the schema. Expert, or well-developed and complex schemas, such as those people hold for themselves, do not change dramatically in response to discrepant information for at least two reasons. First, a well-developed schema will contain many instances of schema-congruent behavior, such that a single inconsistent piece of information carries proportionately little weight. Second, complex schemas have many sub-categories into which discrepant information can be fitted without changing the basic thrust of the schema (Crocker et al. 1984). For instance, if one has a self schema which says that he or she is a very considerate person, there may be subcategories such as "except when interacting with my mother-in-law" and "except when the secretary makes a stupid mistake." Thus, a particular inconsiderate act can be accepted without changing the overall assessment that one is considerate. In a similar vein, Linville (1982) reviews research showing that people whose self concepts are complex and multidimensional make less extreme self-evaluations overall, and are less influenced by a single incident on one dimension than those whose self-concepts are more simple.

Once a self evaluation is formed and integrated with other information in the self schema, it becomes remarkably resistant to change. Several recent studies have documented that self assessments based on false feedback persevere long after debriefing. Subjects led to believe that they have done well (poorly) on a task continue to believe that they are high (low) on the related ability even after being told that the feedback on which the initial evaluation was based was totally unrelated to actual performance or ability (Jennings,
Lepper, & Ross, 1981; Lepper, Ross, & Lau, 1986; Ross, Lepper, & Hubbard, 1975). The perseverance effect is particularly strong when subjects are asked to construct an explanation for their performance before being debriefed (Fleming & Arrowood, 1979). The explanation ties the performance to previous behaviors, dispositions, ability assessments such that the self evaluation remains credible even after the false feedback on which it was originally based is discredited.

It seems likely that superiors have less complex schemas about subordinates' performance than do the subordinates themselves. Less complex schemas tend to either reject inconsistent information all together, especially if other processing demands are high, or to change more dramatically in response to a single piece of salient information (Crocker et al. 1984). If a superior believed that a subordinate was considerate and had a few examples of past considerate behaviors in memory, he or she would probably either ignore an inconsiderate behavior, or change the schema quite suddenly and decide that the subordinate was not really considerate when confronted with an inconsistent behavior.

Recent research by Kozlowski, Kirsch, and Chao (1986) sheds further light on the rating implications of more and less complex schema. They created rater groups who were quite knowledgable or not at all knowledgable about the job on which they would be assessing performance—the job of baseball player. Within each group, each rater selected a player who was familiar and a second player who was unfamiliar, then rated the two players from memory on seven objective performance indices. Raters had previously reported on the "conceptual similarity" of the seven performance dimensions. Conceptual similarity judgments revealed each rater's "implicit performance theory" or schema for the job of baseball player. As predicted, raters knowledgable about the job and
familiar with the ratee showed the least halo and relied less on their conceptual similarity schema. Raters who knew less about the job and/or the specific ratee had to rely more on their implicit theory in the absence of concrete information. The same rationale can explain why superiors' ratings usually contain more halo error than subordinates' self ratings.

At this point it may be helpful to explicitly discuss the role of schemas in producing memory and rating error. At first glance, the literature seems to conflict on these points. On one hand, use of schemas is supposed to enhance recall (Bower & Gilligan, 1979; Kuiper & Rogers, 1979) and make information processing faster and more efficient (Markus, 1977). On the other hand, use of schemas seems to cause a set of typical errors, such as 1) recall of impressions or categorizations rather than the behavior that initially led to categorization, 2) recall of events which may not have happened but which would be consistent with the schema category to which an individual has been assigned, and 3) reliance on the conceptual similarity or illusory correlation among dimensions as specified by the schema rather than on observed covariation of events or behaviors.

A possible resolution to this conflict involves defining three levels or classes of schemas which might be used to process information. The literature has tended to compare only two types at a time, for instance, no schema versus some schema, or simple schema versus expert schema. Considering all three at once helps clarify the puzzling conclusions indentified above. The first level is no schema, or the type of processing that occurs when one is aschematic for the dimension in question. Information is not recalled well, yet neither do false positive memory errors occur.

The second level is the use of a fairly simple schema, such as the one that a superior may have for storing performance information on subordinates. This type of schema should produce the kind of errors mentioned above, because
it is not complex enough to store a great deal of information or make fine distinctions in applying category labels. It will tend to produce extreme judgments and halo error. The third level is a complex or expert schema. This type is sophisticated enough to hold very much information in a variety of categories, and also to be sensitive to actual covariation among events. As Kozlowski et al. (1986) showed, individuals who are highly knowledgable about a job and who have dimensional performance data in memory are able to rate accurately rather than relying on a conceptual similarity schema. The self schema is usually quite complex and well developed, so one would not expect typical schematic errors in self assessment. Admittedly, information input to the self schema may be systematically biased by ego-enhancement errors, but once the information arrives it should be well recalled and carefully used to produce reasonably logical judgments.

Development of Self-Schemas

It appears that work related schemas are developed over time and are elaborated through experience on the job. Lurigio and Carroll (1985) found that parole officers' schemas for categorizing types of offenders were more detailed and consensual among experienced (three or more years) versus inexperienced incumbents. Individuals without relevant job experience at all (clerks in the parole office) were still less schematic regarding types of offenders.

Since the self-schema plays such an important and unique role in assessment, let us now turn to a discussion of how the self-schema for a particular job comes into being. At this point, we really do not know much about how people build up a self-assessment schema for job performance or aspects thereof. However, we speculate that the process might occur as follows. When entering a new performance setting, existing schemas judged as
relevant may be called upon for content. For example, if one has generally high self-esteem and sees him/herself as being competent across a wide range of settings, then the new schema is likely to include the trait, competent. Stone and Stone (1985) found that chronic self esteem does influence self assessments on an unfamiliar task. Other traits on which the performer is strongly schematic may also be seen as relevant. If one is strongly schematic on being a good follower of instructions and being very precise about details, then these dimensions may also occur in one's self-assessment schema for the new task, although they might not be objectively important aspects of the job, or important to one's superior. The performer may then proceed to search for examples of following instructions well and being precise on the job, in order to validate the internally generated hypothesis about performance. In addition to general traits, self-assessment content for a new task is undoubtedly imported from memories (actual or distorted) of past performance on similar tasks or in similar settings.

To this point, we have discussed the "top-down" component of a new schema, which is generated by preexisting schemas (Markus & Sentis, 1982). However, the development of a self assessment schema must also include a bottom-up, or data driven component. The data in this case might be training on what is important and correct in the way of job performance, and performance feedback from a variety of sources. One source is the task itself. In some jobs the success of a behavior is immediately obvious, or can be readily learned by comparing the product to established standards or to the products of others. Other sources of feedback are coworkers, subordinates, superiors, clients, and the formal reward structure of the organization. Greller and Herold (1975) surveyed experienced employees and found that they listed the self and task as more important sources of feedback than coworkers, who were in turn more important than superiors or the organization.
Thus, self-assessment schema dimensions (what is important on this job) and content (examples and judgments of how well I perform on these dimensions) can derive from both internal and external sources. One might suggest that external sources would predominate when an individual is new to a job, especially if the job is quite dissimilar to anything he or she has done before. Thus, newcomers may be quite open to both training on what is important and to feedback on their own performance. However, as experience increases and the self-assessment schema becomes more defined, openness to external feedback may decline. The schema plays a greater role in censoring incoming data, and, as in the Greller and Harold (1975) survey, some sources of feedback may be devalued. This explanation is consistent with Baird’s (1977) and Herold and Parson’s (1980) findings that self and superior assessments of performance diverge as tenure increases.

There is one final factor which may affect the extent to which a complex self schema is developed and used to form performance judgments. This is focus of attention.

Focus of attention has been treated as a disposition, with individuals being chronically high or low in private self-focus (being introspective), and/or high or low in public self-focus (aware of the impression made on others) (Fenigstein, Scheier, & Buss, 1975). A number of other researchers have treated self-focus as a temporary situationally induced state. Private self-focus can be induced by placing the subject in front of a mirror or playing back tapes of the subject’s own voice (Carver & Scheier, 1981). The resulting state is quite similar to what Duval and Wicklund (1972) called “objective self awareness”. Public self-focus can be triggered by placing the subject in front of an observer, a video camera, an obvious one-way window, or an audience, thus raising concern about one’s public impression.
Carver and Scheier (1981, p. 102) believe that self-focus "enhances accessibility of self-schemas." In support of this idea, Turner (1978) found that individuals who were dispositionally high in private self-focus gave significantly longer self-descriptions than those low on this trait. Nasby (1985) also reported that individuals dispositionally high in private self-focus had more fully developed self-schemas. In a recognition task, such subjects displayed the error pattern typical of schema-based processing to a greater extent than individuals low in self-focus.

Carver and Scheier (1981) hypothesize that self-focus increases the tendency to be aware of one's standards or goals, and to more frequently compare one's behavior to standards. A series of four studies (Scheier & Carver, 1983) has shown that individuals high on self-focus (either dispositional or induced) do in fact seek more information about their own performance and choose more diagnostic tasks. The hypothesized comparison to standard is an unobservable cognitive event, but the seeking of information necessary for the comparison seems to indicate that the event occurs. Further, individuals high in self-focus tend to bring their behavior more into line with their standards and attitudes, an outcome which one would expect to follow from more frequent comparisons (Carver & Scheier, 1981). The increased frequency of comparison and awareness of standards should also result in more accurate self-assessment among the highly self-focused. This idea has not been well researched, but one study gives an indication of its promise. Pryor, Gibbons, Wicklund, Fazio, and Hood (1977) found that subjects were more accurate in reporting their past performance when seated in front of a mirror. They asked students to report their SAT scores, and found that inflated reports came from students whose actual scores were below the median and who were not self focused. Low scoring students who responded in the presence of a mirror were much more accurate.
Clearly, self-focus is something which will not affect a superior’s rating of a subordinate, but which may color the subordinate’s self perception. Individuals dispositionally high on private self-focus may have more detailed and accurate self-schema, may compare themselves to goals and standards more often, and in consequence may be able to assess themselves quite realistically. On the other hand, individuals high on public self-focus may compile a great deal of biased information on their performance as they strive to manage the impressions of others towards a positive assessment. Finally, individuals chronically low on both types of self-focus may have a paucity of information on themselves, or may not bother to access the information they do have and may not compare to standards very frequently. Thus, their self-assessments will probably be inaccurate, but may be as likely to be too low as too high. A summary of schema theory and implications appears in Table 5.

Other Influences on Superior/Subordinate Agreement

This paper is not intended to provide a comprehensive picture of superior cognitive processes in rating--that has been done by others. The purpose instead is to highlight the self rating process and contrast it to the superior rating process where the two seem to diverge. To do this thoroughly, two further topics require discussion: the information environment in which performance information is gathered, and the motivation of the superior as a rater.

The Information Environment

The amount and diagnosticity of information available to assess performance may be quite different for performers as opposed to superiors. Amount of information should vary greatly for several reasons. The first is
Table 5

Schema Theories

<table>
<thead>
<tr>
<th>Theory</th>
<th>Implications</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schemas provide a framework for storing, interpreting, and recalling</td>
<td>Subordinates will have complex self schemas with which to evaluate their</td>
<td>Superior/subordinate assessments disagree at least partly because each</td>
</tr>
<tr>
<td>information.</td>
<td>performance. They will have much more performance relevant information in</td>
<td>party is applying a different rating schema, because subordinates have much</td>
</tr>
<tr>
<td>Schema-based processing may result in</td>
<td>memory than do their superiors.</td>
<td>more information available in memory, and because superiors tend to judge</td>
</tr>
<tr>
<td>particular types of errors:</td>
<td>Superior and subordinates have different (but consistent within type of</td>
<td>others on the dimensions which are most relevant for judging themselves.</td>
</tr>
<tr>
<td>Recall of impressions or category membership rather than the concrete</td>
<td>rating schemas, as shown by strong source-of-rating factors in multi-trait,</td>
<td></td>
</tr>
<tr>
<td>instance of behavior which led to the impression or categorization.</td>
<td>multi-rater studies.</td>
<td></td>
</tr>
<tr>
<td>(supported)</td>
<td>Superior's less complex schemas will be more prone to typical errors than</td>
<td></td>
</tr>
<tr>
<td>Recall of events which did not occur but which are typical of persons</td>
<td>will than subordinates' self ratings.</td>
<td></td>
</tr>
<tr>
<td>in that category. (Supported)</td>
<td>(supported)</td>
<td></td>
</tr>
<tr>
<td>Illusory correlation between events or traits assumed to co-occur.</td>
<td>Subordinates will tend to evaluate themselves on the dimensions on which</td>
<td></td>
</tr>
<tr>
<td>(supported)</td>
<td>they are strongly schematic, whereas superiors will apply their personally</td>
<td></td>
</tr>
<tr>
<td>Individuals are schematic on dimensions which they consider important</td>
<td>most important dimensions to those they rate.</td>
<td></td>
</tr>
<tr>
<td>or on which they are extreme. These dimensions are used to evaluate</td>
<td>Superior who have previously performed a subordinate's job may develop an</td>
<td></td>
</tr>
<tr>
<td>both the self and others. (supported)</td>
<td>evaluation schema for the job similar to the schema used by the subordinate,</td>
<td></td>
</tr>
<tr>
<td>Schemas allow one to process schema-relevant information more quickly,</td>
<td>thus enhancing potential agreement.</td>
<td></td>
</tr>
<tr>
<td>to recall more schema-relevant information, and to resist inconsistent</td>
<td>Self assessments may be grounded in a network of supporting memories going</td>
<td></td>
</tr>
<tr>
<td>information. (supported)</td>
<td>back many years rather than in present performance. Superior assessments are</td>
<td></td>
</tr>
<tr>
<td>Self schemas are very complex structures, able to store a great deal</td>
<td>grounded largely in present performance, so the potential for disagreement is</td>
<td></td>
</tr>
<tr>
<td>of information. (supported)</td>
<td>high.</td>
<td></td>
</tr>
<tr>
<td>Self assessments persevere even in the face of contradictory evidence.</td>
<td>Some individuals are dispositionally likely to develop more complex self</td>
<td></td>
</tr>
<tr>
<td>(supported)</td>
<td>schemas and to access them often. These individuals may be able to give</td>
<td></td>
</tr>
<tr>
<td>Dispositional and situational focus of attention affects the frequency</td>
<td>fairly accurate self assessments.</td>
<td></td>
</tr>
<tr>
<td>which behavior is compared to standards or schemas.</td>
<td>Individuals give more accurate self assessments when high self-focus is</td>
<td></td>
</tr>
<tr>
<td></td>
<td>situationally induced.</td>
<td></td>
</tr>
</tbody>
</table>
opportunity to observe behavior. The performer necessarily observes all of his or her own work behavior whether it occurs in or out of the office, during regular working hours or nights and weekends. The superior is limited to observing only during the regular work day. Further, most superiors do not have the subordinate constantly in view, and even if they do, cannot pay continuous attention to the actions of a single subordinate. (Of course, real-time observation is not the only source of information, as raters may also be able to evaluate work products or outcomes.) Second, subordinates may very actively seek additional information on their performance. Ashford and Cummings (1983) suggest that incumbents engage in feedback seeking behaviors such as asking for feedback from various sources, and attentively monitoring others for subtle clues as to how one's performance is being perceived. Thus, subordinates acquire a great deal more performance information on themselves than do their superiors.

As discussed earlier, self-schemas tend to be complex and well developed, and thus capable of storing a great deal of information. The superior will not have as fully developed a schema for storing information about subordinate performance, and this schema or set of schemas must usually store information on more than one subordinate. Thus, when it is time to make a performance judgement, the superior will have both observed less behavior and recalled less of it than the performer.

The diagnosticity of the information possessed by the two parties may also differ. Superiors as observers do not have full access to internal cognitions of the actor concerning intent, overall performance strategy, intrinsic satisfaction, effort level, or the like. Recent research by Andersen and her colleagues has shown that private thoughts and feelings are considered much more diagnostic and informative of what one is really like
than are overt behaviors (Andersen and Ross, 1984). Both actors and observers agreed that thoughts and feelings were most helpful, and observers were actually able to form objectively more accurate social impressions following exposure to this type of information than exposure to reports of observable behavior (Andersen, 1984).

Further, Andersen and Williams (1985) have shown that individuals do make use of the thoughts and feelings not usually available to observers when they engage in private self-evaluation. Specifically, subjects instructed to privately recall their positive thoughts and feelings about past events subsequently raised their self-esteem more than individuals who recalled their positive behaviors during the events. Thus, in making performance judgments, actors not only have access to information with other raters do not have, but they also use this highly diagnostic information in self-evaluation. At the same time, superiors may be seeking and using a different type of information which they consider highly diagnostic. DeNisi et al. (1984) suggest that evaluators' tendency to seek and overweight negative information (Bolster & Springbett, 1961) flows from the relative rarity and thus high distinctiveness of this type of information. Fisher's (1979) finding that supervisors of poor performers were able to rate and give feedback after fewer trials than supervisors of high performers is consistent with this explanation. The superior's greater sensitivity to negative information contrasts markedly with the subordinate's capability to ignore such unfavorable information.

Rater Motivation and Purpose of Appraisal

The motivation of the self as a rater has already been discussed, and evidence of motivation toward accurate self assessment, ego-enhancing
self-assessment, and self-presentational concerns has been presented. Some aspects of superior rater motivation have also been mentioned, but additional discussion is needed.

The recent literature has included a number of calls for studying not just the ability of raters to make sound performance judgments, but also their willingness to record these judgments accurately under field conditions (Banks & Murphy, 1985; DeCotiis & Petit, 1978; Landy & Farr, 1980; Wexley & Klimoski, 1984). There is evidence that raters often are not motivated to provide accurate ratings. DeCotiis and Petit note that the consequences of rating accurately are often negative. Fisher's (1979) raters said they feared that ratees would no longer like them if below average ratings were made and feedback face-to-face. When Bernardin and Villanova (1984) asked supervisors the extent to which 20 factors contributed to rating inaccuracy, the highest rated factor (3.8 on a five point scale) was, "Raters rate higher than deserved because they prefer to avoid confrontations."

The studies of purpose of rating tend to verify these conclusions. Raters are more lenient when they expect to personally give feedback and explain their ratings (Fisher, 1979; Sharon & Bartlett, 1969) or when the ratings could have a negative impact upon the employment status of the ratees (Gallagher, 1978; McIntyre, Smith, & Hassett, 1984; Zedeck & Cascio, 1982). Ratings made in confidence for research purposes are typically less lenient.

It is not yet clear how much of the leniency phenomenon is intentional distortion, as admitted to by Bernardin and Villanova's (1984) respondents, and how much may be unintentional. There is some evidence that different purposes automatically trigger somewhat different schema or standards and produce legitimately different judgments (DeNisi and Williams, 1986; Williams, DeNisi, Blencoe, and Cafferty, 1985). Zedeck and Cascio (1982) found that
raters used different weights when combining information on five dimensions into an overall assessment when the purpose of ratings was to award a merit raise versus to decide on discharge. Thus, when making an objectively defined leniency error, raters may or may not be aware that their rating is higher than it should be.

Given that subordinates usually rate themselves more highly than do their superiors, any leniency tendency on the part of superiors should decrease disagreement between the two rating sources. This might be expected to make the feedback and performance discussion steps easier and less confrontative (though perhaps less productive), especially if superiors are unaware that their ratings are lenient. However, if the superior has intentionally inflated the ratings and feels that he or she already has given the subordinate "a break", then the superior may react quite negatively and inflexibly to any remaining disagreement over performance level. This section is summarized in Table 6.

Summary Model and Suggestions for Increasing Agreement

Figure 3 depicts some of the differences in the performance judgment processes used by superiors and subordinates that have been suggested in this paper. The differences between superior and subordinate performance assessment processes will be summarized below, working from the beginning stages of perceiving information, through storing and retrieving it, and finally to making a performance judgment.

Both parties begin by drawing information from the environment. Superiors have less time and attention to observe performance relevant behavior than do subordinates. However, superiors may also have information
Table 6

Other Influences on Agreement

<table>
<thead>
<tr>
<th>Research</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subordinates have greater opportunity to observe their own performance than do superiors.</td>
<td>Subordinates, being better informed, may be able to make more accurate judgments than superiors.</td>
</tr>
<tr>
<td>Subordinates actively seek performance information.</td>
<td>This tendency may increase agreement between superiors and subordinates.</td>
</tr>
<tr>
<td>Subordinates have access to internal cognitions relevant to performance, while superiors do not. Internal information is highly diagnostic in making accurate assessments. (supported)</td>
<td>Superior ratings made for some purposes may agree more with self ratings than ratings made for other purposes.</td>
</tr>
<tr>
<td>Raters are motivated to avoid giving accurate negative evaluations, particularly when such ratings must be feedback or will affect the subordinates’ conditions of employment.</td>
<td>(supported)</td>
</tr>
</tbody>
</table>
Figure 3

Reconstruction of Events
Based on Category Membership
Use of Self as a Comparison
Reluctance to Communicate
Negative Feedback

Information Perceived
- Information Unique to Superior
- Performance Information Environment
- Internal Performance Cues
- Feedback Seeking Behaviors

Information Perceived
- Differential Opportunity to Observe

Information Stored and Recalled in Schema for Subordinate Performance
- Actor-Observer Bias
- Differences in Schema
  1. Complexity
  2. Information Capacity
  3. Dimensions

Information Stored and Recalled in Self-Schema
- Consistency Bias
- Resistance of Complex Schema to Disconfirmation
- Perseverance Effects
- Accuracy vs Ego-enhancement Motives
- Self-focus

Self Assessment
- Choice of Comparison Other
- Self Presentational Concerns

Superior Assessment
not directly available to subordinates, such as records of complaints, inspector rejections, or higher ups' reactions to subordinates' work. Subordinates have a larger information environment, in that they also have access to internal thoughts concerning previous performance in other settings, intentions, and effort level which the superior does not have. Further, subordinates may actively seek feedback from a variety of sources, thus increasing the amount of performance relevant information they possess. Not all of the available performance information is entered into long term memory. For both parties, probably only information which is considered relevant or diagnostic is remembered. Thus, performance events which the perceiver thinks were not under the control of the actor will be disregarded. As mentioned earlier, subordinates and superiors may disagree when making these attributions, and so may recall as relevant somewhat different sets of performance information. Further, each party will tend to accept as credible only information which is largely consistent with their existing evaluations.

The mechanism controlling memory is the schema. I have suggested that the self-schema is more complex than the superior's performance rating schema, so the subordinate has a much greater storage capacity for incidents of performance. Differences in schema complexity also determine the way in which discrepant information is handled. Further, superiors and subordinates may differ in the performance dimensions considered relevant, with each individual seeing as most important those dimensions on which he or she excels. For the subordinate, level of self-focus (either dispositional or situational) probably affects the degree to which performance information is perceived, remembered, and compared to the standard provided by the schema.

The motivation of the parties also can affect the information gathering, storage, and judgment process. There is evidence that purpose of observation affects the way events are encoded in memory, and that purpose of rating is
related to leniency/severity of rating. It is not clear whether the more or less severe ratings given under different purposes reflect true differences in the superior's actual opinions, or conscious distortions in order to spare the subordinate or ease the process of giving feedback.

On the subordinate side, the motivation to self assess accurately may conflict with the motivation to see oneself positively. The motive which is operative in a given situation may affect what information is stored and processed, and also what standard or "other" is chosen for comparison purposes.

It seems likely that some disagreement could occur if the superior chooses a different comparison other than the subordinate, specifically, the superior's own real or imagined performance on the subordinate's job. As well as being a standard with which the subordinate is unfamiliar and unlikely to use for self-assessment, this practice also is likely to activate a host of ego-defensive mechanisms in the superior.

Thus, there are a number of reasons why superiors and subordinates may reach different conclusions about the subordinate's performance. Each party is subject to its own set of cognitive and motivational biases. It should not be surprising that superior-subordinate agreement is low when the two are drawing from somewhat different information environments, processing information through different schemas, and making judgments with different motives. Clearly, some disagreement is "legitimate". It is equally clear that neither party's judgment is likely to be particularly accurate or correct.

Increasing Agreement

The model suggests several possibilities for improving agreement. In order for the final evaluations to agree, it would seem that both the information input and processing steps would need to be similar. To increase the similarity of information input, superiors and subordinates should discuss and evaluate each performance event as it occurs. They should attempt to reach
A consensus on the causes underlying the event and the goodness/badness of the particular performance. Both parties can use these sessions to share information to which the others would not normally have access. Thus, after each performance event, both would store essentially the same evaluation of that event in memory. Given the failings of memory, and the censoring function of schemas, it might be wise for both parties to maintain a diary of these performance events. Open discussion of events as they occur would also allow the superior to communicate his or her dimensional performance schema to the subordinate, and for a common set of dimensions on which both agree to emerge.

A more formal approach would feature rater training. There has been a great deal of recent research on training superiors to rate with less error and more accuracy (Smith, 1986), but no mention of providing similar training to the subordinates who will be evaluated. The prevailing practice of training only superiors may actually increase superior/subordinate disagreement, because subordinates typically display a great deal of leniency error and reduced leniency is often a product of rater training for superiors. Both superiors and subordinates should be trained so that a common performance schema can be adopted by each.

Further, such training should occur before performance rather than after performance but before rating (the usual timing of rater training). Information is processed and stored according to the schemas in existence at the time the performance is observed, so it is desirable for superiors and subordinates to have similar schemas at that time (Williams et al., 1985). Attempting to later impose a common schema on information processed and stored under differing schemas will probably be ineffective. Finally, rater training should be expanded in content, so that both parties can be made aware of biases that they are likely to display in attributions, in using the self as a referent, and in applying their personally most relevant dimensions to others.


Manpower R&D Program - List A

(One copy to each addressee except as otherwise noted)

Director Technology Programs
Office of Naval Research (Code 12)
Arlington, VA 22217-5000

Director Research Programs
Office of Naval Research (Code 11)
Arlington, VA 22217-5000

Chairman, Manpower R&D Planning Committee
Office of the Chief of Naval Research
Code 222
Arlington, VA 22217-5000

Life Sciences Technology Program Manager
Office of Naval Research (Code 125)
Arlington, VA 22217-5000

Mathematics Programs
Office of Naval Research (Code 1111MA)
Arlington, VA 22217-5000

Head Information Sciences Division
Office of Naval Research (Code 1133)
Arlington, VA 22217-5000

Associate Director for Life Sciences
Office of Naval Research (Code 114)
Arlington, VA 22217-5000

Head, Psychological Sciences
Office of Naval Research (Code 1142)
Arlington, VA 22217-5000

Engineering Psychology Programs
Office of Naval Research (Code 1142EP)
Arlington, VA 22217-5000

Group Psychology Programs
Office of Naval Research (Code 1142CF)
Arlington, VA 22217-5000

Personnel and Training Research Programs
Office of Naval Research (Code 1142PT)
Arlington, VA 22217-5000

Defense Technical Information Center
(12 copies)*
DTIC/DDA-2
Cameron Station, Building 5
Alexandria, VA 22314

*If report is ready for unlimited public distribution
Dr. Lawrence R. James
School of Psychology
Georgia Institute of Technology
Atlanta, GA 30332

Dr. John R. Frederiksen
Bojt Beraneck & Newman Inc.
50 Moulton Street
Cambridge, MA 02238

Dr. Lee Roy Beach
Department of Psychology (N1-25)
University of Washington
Seattle, WA 98195

Dr. Cynthia D. Fisher
College of Business Administration
Texas A&M University
College Park, TX 77843

Dr. Barry Riegelhaupt
Human Resources Research Organization
1100 South Washington Street
Alexandria, VA 22314

Dr. Lawrence Goldberg
Economic Research Laboratory
1914 Association Drive
Reston, VA 22091

Dr. James C. Coyne
Mental Research Institute
555 Middlefield Road
Palo Alto, CA 94301

Dr. Robert P. Archer
Eastern Virginia Medical Authority
P. O. Box 1980
Norfolk, VA 23501

Dr. Meg Gerrard
Psychology Department
Iowa State University
Ames, Iowa 50010

Dr. Jomille H. Braddock II
Center for the Social Organization
of Schools
The Johns Hopkins University
3505 North Charles Street
Baltimore, MD 21218

Dr. Eric Flamholtz
Graduate School of Management
UCLA
Los Angeles, CA 90024

Dr. Richard C. Morey
Graduate School of Business Admin.
Duke University
Durham, NC 27706

Library
Naval Training Systems Center
Orlando, FL 32813

Library
Naval War College
Newport, RI 02940

Chief, Survey and Market
Analysis Division
Defense Manpower Data Center
1600 Wilson Boulevard, #400
Arlington, VA 22209
Director, Human Resource Management
Training Department
Naval Amphibious School
NAB Coronado, CA 92155

Commanding Officer
Organizational Effectiveness Center
Naval Training Center Building 304
San Diego, CA 92133

Commanding Officer
Organizational Effectiveness Center
Pearl Harbor, HI 96860

CINCPACFLT
Organizational Effectiveness Division
Code 71
Pearl Harbor, HI 96860

Commanding Officer
Organizational Effectiveness Center
1211 South Fern Street, Room A112
Arlington, VA 22202-2808

Commanding Officer
Organizational Effectiveness Center
5621-23 Tidewater Drive
Norfolk, VA 23509

Commander in Chief, U.S. Atlantic Fleet
Organizational Effectiveness Division
Code NO41
Norfolk, VA 23511

Director, Human Resource Management
Training Department
Naval Amphibious School
NAB Little Creek
Norfolk, VA 23521
END
12-86
DTIC