Egoistic and Moralistic Biases in Self-Perception: The Interplay of Self-Deceptive Styles With Basic Traits and Motives

Delroy L. Paulhus
University of British Columbia

Oliver P. John
University of California, Berkeley

ABSTRACT The literature on personality traits and defense mechanisms suggests individual differences in two self-favoring tendencies, which we label “egoistic bias” and “moralistic bias.” The two biases are self-deceptive in nature and can be traced to two fundamental values, agency and communion, that impel two corresponding motives, nPower and nApproval. The two sequences of values, motives, and biases form two personality constellations, Alpha and Gamma. Associated with Alpha is an egoistic bias, a self-deceptive tendency to exaggerate one’s social and intellectual status. This tendency leads to...
unrealistically positive self-perceptions on such traits as dominance, fearlessness, emotional stability, intellect, and creativity. Self-perceptions of high Alpha scorers have a narcissistic, “superhero” quality. Associated with Gamma is a moralistic bias, a self-deceptive tendency to deny socially deviant impulses and to claim sanctimonious “saint-like” attributes. This tendency is played out in overly positive self-perceptions on such traits as agreeableness, dutifulness, and restraint. The Alpha-Gamma conception provides an integrative framework for a number of central issues in personality psychology.

It is not surprising that many writers and researchers have attempted to link defense mechanisms to major personality traits. After all, one of the fundamental assumptions in psychodynamic theory is that defensive processes play a major role in character formation (e.g., Cramer, 1991; Haan, 1977; Shapiro, 1981). And it follows that some, if not all, important personality traits are influenced by defense mechanisms (for a recent review, see Paulhus, Fridhandler, & Hayes, 1997).

**Defenses and Dimensions of Personality: A Selective Review**

Our review will focus on the few research programs that have been at once influential and empirical in nature. To organize our discussion of personality dimensions, we have adopted the Five-Factor Model (or “Big Five”) as our framework of choice. Our labels for the five factors are consistent with the consensus reached in a recent special issue of this journal (McCrae & John, 1992), namely, Extraversion, Agreeableness, Conscientiousness, Emotional Stability (vs. Neuroticism), and Openness to Experience.

**Early MMPI Work**

The interplay between defense and personality was evident in early structural analyses of the MMPI—for many years the most well-tilled bed of personality items. The clinical scales were developed via the contrasted groups method and were targeted, of course, at clinical syndromes. When administered to normal samples and factor-analyzed, however, the MMPI-1 item pool also yielded a set of “content scales” covering a wide array of normal personality variables (Wiggins, 1966).
The first content scale, based on the largest factor, was usually interpreted as an overall index of maladjustment, distress, or neuroticism (Block, 1965; Johnson, Null, Butcher, & Johnson, 1984). Its surprisingly high correlation with ratings of social desirability, however, led some to believe that the first MMPI factor represented, for the most part, response style contamination (e.g., Edwards, 1957; Jackson & Messick, 1958): that is, low scorers were willing to admit negative information about themselves whereas high scorers were “defensive” and simply gave the desirable answer. To outside observers, this very public disagreement was an unsettling indictment of the value of personality assessment: The whole approach seemed dubious because measurement specialists could not even agree among themselves whether the first factor of personality represented substantive personality differences or some kind of defensive process.¹

Byrne’s Repression-Sensitization

The widely used Repression-Sensitization (R-S) scale (Byrne, 1961) was an early attempt to measure the defensive personality (“the repressor”) as a trait. Developed from items showing maximum contrast between groups of clinically rated “defensives” and normals, the R-S scale constituted a rather motley parade of item content (e.g., “Criticism or scolding hurts me terribly”; “I almost never dream”). After its refinement, however, the R-S overlapped considerably with measures of trait anxiety. And neither was empirically distinguishable from the first factor of the MMPI (Liberty, Lunneborg, & Atkinson, 1964). The final verdict was that, although it may tap repressive style, the R-S scale is inextricably confounded with trait anxiety (Bonanno & Singer, 1990). In short, the history of the R-S scale recapitulated the earlier debate about whether the first factor of the MMPI was substantive (i.e., differentiating high from low anxiety) or defensive (i.e., differentiating repressive from expressive individuals). It combined features of both.

¹. This particular factor was largest because of the selective content of the MMPI, of course. In inventories with more representative samples of items, all of the Big Five loaded on the desirability factor (Saucier, 1994).
Crowne and Marlowe’s Approval Motive

Certain that defensive style could be measured separately from psychopathology, Crowne and Marlowe (1960) developed a set of items that referred instead to everyday personal and interpersonal behaviors (e.g., “I’m always willing to admit it when I make a mistake”). The resulting Marlowe-Crowne (MC) scale was indeed relatively independent of the first factor of the MMPI.

Rather than simply being an index of questionnaire response style, the MC scale showed a surprising degree of temporal stability and a clear set of behavioral correlates: In short, it appeared to be a trait in its own right (Crowne & Marlowe, 1964). Accordingly, agreement with the MC items came to be viewed as a symptom of the approval-seeking character.² Such characters were congenial, conforming, and cautious, although their self-descriptions seemed to be remarkably free of sin.

The interpretation of the MC as approval-seeking can even explain why high scorers can be induced to cheat and lie if it is necessary to avoid disapprobation from authority figures (Millham & Jacobson, 1978). Approval-seeking is also consistent with research demonstrating that MC scores drop considerably if subjects are questioned under bogus pipeline conditions as opposed to standard conditions (Derakshan & Eysenck, in press). Developmentally, high scorers on the MC were said to have acquired a sensitivity to parental approval in childhood. Subsequently, therefore, their adult behavior is oriented toward avoidance of disapproval by being socially normative (Crowne, 1979). Here was a classic case of the close interplay between defense and personality.

Most relevant to the present article is McCrae and Costa’s (1983) claim that MC scores represent accurate reports of one’s true social desirability. Their argument is based on the finding that self-report MC scores correlate substantially with ratings of Agreeableness and Conscientiousness by knowledgeable peers. Hence, high MC scorers really do have the agreeable and conscientious personalities that win social approval.³ McCrae and Costa paid less attention, however, to the fact that the MC correlated much more highly with self-reports of Agreeableness and Conscientiousness than with peer reports of those same traits. In other

² Subsequent research suggested that a more accurate interpretation was avoidance of disapproval (Crowne, 1979).
³ This interpretation as “socially desirable character” makes it difficult to explain their cheating and lying (Millham & Jacobson, 1978).
words, the MC captures some substance (as McCrae and Costa emphasize) but also includes self-report method variance on Agreeableness and Conscientiousness that departs from reality.

**Weinberger’s Repressive Style**

Weinberger and colleagues have taken a different approach to the ambiguity over the first factor of the MMPI (e.g., Weinberger, Schwartz, & Davidson, 1979). They agreed that a fundamental kind of defensiveness, worthy of the label “repression,” could be attributed to those with low scores on measures of distress (trait anxiety)—but only if they also scored high on a measure of restraint. This fertile typological operationalization yielded a broad array of correlates consistent with the concept of repression (Weinberger, this issue).

For more details, we refer the reader to Weinberger’s contribution and simply note two points relevant to our article. First, Weinberger and colleagues have highlighted the importance of the restraint dimension by naming it as one of the two factors underlying psychopathology and linking it to defense. Second, some of the research in the Weinberger tradition has revealed defensive correlates (e.g., symptom denial) for the Distress and Restraint factors as individual dimensions, rather than only in combination (e.g., Bonanno, Siddique, Keltner, & Horowitz, 1997; Jamner & Schwartz, 1987; Warrenburg et al., 1989).

**Self-Deception and Impression Management**

A competing approach to the role of defense in personality began with the goal of separating conscious bias from unconscious bias (Sackeim & Gur, 1978). This distinction motivated the development of a pair of complementary measures of socially desirable responding (Paulhus, 1984). First was the Impression Management scale, a measure designed to tap purposeful tailoring of responses to impress an audience. Second was the Self-Deceptive Enhancement scale, a measure designed to tap an unconscious favorability bias in self-descriptions. Together these two measures formed the Balanced Inventory of Desirable Responding, which went through several revisions (Paulhus, 1984, 1988, 1998b). It was recommended that the effects of impression management on self-reports be controlled whereas self-deception should not be controlled because of its direct role in personality traits (e.g., Paulhus, 1986, 1991).
The Self-Deceptive Enhancement scale showed clear evidence of uncon- 
scious motivated distortion (Hoorens, 1995; Paulhus, 1988, 1998b), as 
well as adaptive value (Bonanno et al., 1997; Jamner & Schwartz, 1987; 
Paulhus, 1998a). In contrast, scores on the Impression Management scale 
showed a special sensitivity to situational demands, for example, public 
versus anonymous administration conditions (Paulhus, 1984, 1988).

Under relatively anonymous conditions, the two measures of socially 
desirable responding show a distinctive pattern of correlates with the Big 
Five personality traits (Paulhus, 1988; Meston, Heiman, Trapnell, & 
with Extraversion and Openness to Experience, and to a lesser extent 
with Emotional Stability. In contrast, the Impression Management 
scale correlates most strongly and positively with Agreeableness and 
Conscientiousness.

**Narcissism**

Of long-standing interest to clinicians, the concept of narcissism has 
recently received close attention as a normal personality characteristic 
(e.g., Cramer, 1995; Emmons, 1987; John & Robins, 1994; Raskin & 
Hall, 1981). According to the DSM-IV criteria, the defining charac-
teristics of narcissism include a grandiose sense of self-importance, a 
tendency to exaggerate accomplishments and talents, and an expectation 
to be noticed as “special” even without appropriate achievement. All 
clinical accounts of narcissism (e.g., Kohut, 1971; Millon, 1990) concur 
that narcissistic individuals hold unrealistically exaggerated beliefs about 
their abilities and achievements. In support of these accounts, research 
on nonclinical populations has shown that more narcissistic individuals 
respond to threats to their self-worth by promoting and exaggerating their 
assets (Raskin, Novacek, & Hogan, 1991a; John & Robins, 1994; Robins 
& John, 1997a, 1997b; Paulhus, 1998a) and by denigrating others (Morf 
& Rhodewalt, 1993).

Although these experimental studies clearly demonstrate the strong 
self-favoring bias of narcissistic individuals, this bias appears only on 
select traits (e.g., Extraversion, Openness, Intelligence), not on other 
socially approved traits such as Agreeableness or Conscientiousness 
(Paulhus & John, 1994). There is little doubt that narcissists actually 
believe their claims to superiority (Robins & John, 1997b), consistent 
with a self-deceptive bias. Finally, narcissism is substantially and positively
correlated with self-deception but not with forms of desirable responding relating to social conventionality, for example, the Marlowe-Crowne or Impression Management scales (Paulhus, 1998b; Raskin, Novacek, & Hogan, 1991b).

How is narcissism related to the major personality traits? When measured by the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1981), narcissism can be located at the Arrogant-Calculating position within the two-dimensional interpersonal circumplex (Leary, 1957; Wiggins & Pincus, 1994) or in the disagreeable extravert location in Big Five personality space (Widiger, Trull, Clarkin, Sanderson, & Costa, 1994). 4

Other Defenses

There is little research linking larger batteries of defense measures directly to the Big Five personality traits (see Paulhus et al., 1997). A notable exception is a recent study by Costa, Zonderman, and McCrae (1991) that provides Big Five correlates of (1) Haan’s (1977) MMPI defense and coping measures, (2) Gleser and Ihilevich’s (1969) Defense Mechanisms Inventory (DMI), and (3) Bond, Gardner, Christian, and Sigal’s (1983) Defensive Style Questionnaire (DSQ).

Among DSQ correlates, it was found that individuals high in Agreeableness use self-sacrificing defenses, whereas use of the more adaptive (esteem-promoting) defenses was common among individuals high on Extraversion and Openness, but low on Neuroticism. On Haan’s MMPI defense scales, high Denial individuals also scored lower on self-reports (but not peer reports) of Neuroticism. Overall, Costa and colleagues (1991) concluded that longitudinal studies provide support for the stability of defenses and that there is cross-observer evidence that defenses are related to traits.

Coping Styles

Coping styles have traditionally been distinguished from defense mechanisms (e.g., Haan, 1977). Yet this distinction has recently become less clear-cut, with a number of writers arguing for an integration of the two kinds of mechanisms (for a review, see Paulhus et al., 1997). As with the

4. The clinical version, combining arrogance with insecurity, is more difficult to position in normal personality space.
R-S scale, some critics have argued that coping measures are hopelessly confounded with neuroticism (e.g., Skodal, Dohrenwend, Link, & Shront, 1990). McCrae and Costa (1986) also found associations with Neuroticism, as well as with other Big Five dimensions, but found these associations to constitute meaningful trait-initiated coping choices.

More recently, O’Bien and DeLongis (1996) framed these associations even more positively by demonstrating the utility of trait concepts for organizing the plethora of coping measures in the literature. First, the various coping measures could be usefully organized in terms of agentic and communal strategies. Second, coping outcomes could be understood in terms of interactions between Big Five personality traits and situational factors.

**Dimensions of Self-Favoring Bias**

Some recent work conducted by the present authors (Paulhus & John, 1994) introduced a novel methodology to explore the relation between self-deceptive mechanisms and fundamental personality traits. Our methodology was predicated on John and Robins’s (1994) notion that a self-deceptive bias can be indexed by systematic departures from social reality. But how can actual personality differences be differentiated from bias?

Measuring such departures from social reality required a new unit of analysis: the self-criterion residual (SCR). Rather than difference scores (e.g., Colvin, Block, & Funder, 1995; Wylie, 1961), the bias index is calculated as the residual variance that remains when a self-report measure is regressed on a corresponding criterion measure of that same variable. Thus, all the valid criterion variance is removed from the self-reports. For example, scores on a self-report Agreeableness scale are regressed on a set of peer ratings of the same items and the self-report residual is isolated as a separate variable. Because all the self-report variance shared with the peer ratings has been removed, the residual represents only self-report inflation: High scores indicate overclaiming agreeableness (and low scores indicate underclaiming) relative to the peer-rating criterion. These self-favoring bias scores can then be correlated with any other variable in the data set.

5. For a recent discussion of the “criterion problem” in research on accuracy and bias, see Robins and John (1997a).
In preliminary work using the SCR approach, we conducted a series of factor analyses on a comprehensive set of personality and intelligence bias scores (Paulhus & John, 1994). The data sets were two large samples for which we had obtained both self-ratings and ratings by well-informed others on nine traits covering the Big Five domains. For each trait, we calculated the residual score by indexing the degree to which the self-rating exceeded the peer-rating criterion. This procedure yielded nine self-criterion bias scores (i.e., one for each trait), which we then intercorrelated. A factor analysis of this matrix showed that there were two, rather than five, self-favorability bias factors: one was marked by self-favoring on Agreeableness, Dutifulness, and Nurturance and the other by self-favoring on Extraversion, Dominance, Intellect, and Openness (see Figure 1). Similar factors were found in a Chinese sample from Hong Kong (Yik, Bond, & Paulhus, 1998).

**ALPHA-bias**
- Dominance
  - Extraversion
  - Intellect

- Neuroticism
  - Ambition

- Dutifulness
  - Nurturance

**GAMMA-bias**
- Agreeableness

---

6. Neuroticism is reverse-scored and labeled “Stability” to ensure that all domains point in the positive direction.

7. We found it useful to separate Conscientiousness into Dutifulness and Ambition following Paunonen and Jackson (1996). For similar distinctions, see McCrae and John (1992) and Wiggins and Trapnell (1996).
These two forms of self-favoring bias seemed to carry connotations of “Saints” and “Superheroes.” The first factor captures a tendency to see oneself as more altruistic and dependable than would seem justified by others’ judgments: These individuals are overclaiming possession of traits indicating moral virtue and respect for social convention. The second factor captures a tendency to see oneself as more competent socially and intellectually than one is seen by well-informed others. Such individuals are claiming superiority over others by overclaiming possession of traits associated with competence and social status.

Note the extraordinary flexibility of this SCR methodology. Any self-rated trait that has a credible criterion can be included. The method is potentially comprehensive in permitting the simultaneous analysis of discrepancies for a broad range of individual differences. In a sense, the approach represents the analysis of self-report method variance, that is, that part of self-reported personality that is unique to the self-report technique.

**Crystallizing Two Themes**

In sum, the literature reviewed above reflects a variety of forms of favorability bias in self-perception and self-reports. In our view, however, the various versions of self-favorability cluster into two themes. The first is an egoistic tendency to see oneself as exceptionally talented and socially prominent; the second is a moralistic tendency to see oneself as an exceptionally good member of society.

How are these two tendencies related to traditional defense mechanisms, defined here (Cramer & Davidson, this volume) as unconscious mental operations that minimize anxiety or maintain self-esteem? Certainly, the two go far beyond “response styles,” that is, tendencies to bias one’s responses on questionnaires (Messick & Jackson, 1958; Paulhus, 1991). Although they show some departure from reality, the biased self-descriptions appear to be held honestly—even on totally anonymous reports. At the same time, we doubt that the two correspond directly to any two traditional defense mechanisms. Instead, we allude to Sackeim’s (1983) argument that self-deception is the common ingredient in defense mechanisms. Following that argument, we conceive the processes responsible for the two self-favorability biases as self-deceptive styles.

Sackeim’s (1983) influential article proposed that the efficacy of a defense mechanism rests on the interaction of two mutually conflicting
representations of the same information (see also Sackeim, 1988): that is, some representation of the threatening information is required to effectively avoid or distort it. But unawareness of one representation is also a necessary condition (p. 245). In short, the distortion process involves exploiting information about the discrepancy without acknowledging it. Evidence that such self-deception can occur was established in the widely cited article by Gur and Sackeim (1979).

Our present arguments derive from a parallel discrepancy between two social realities—one defined by the self and one defined by informed observers. Rather than verifying that the self is aware (at some level) of the observer perspective, we study social information that should be continually available to the self, namely, basic traits. Two forms of such departures from reality were isolated in our research on self-criterion residuals detailed earlier: two tendencies to believe that one possesses positive qualities that are not corroborated by close acquaintances. Below we present further evidence that these biases are self-deceptive in nature and broad in scope.

**Assembling the Constellations of Alpha and Gamma**

After examining related literatures, we became convinced that the two self-deceptive mechanisms identified above were components of two larger constellations of constructs linked with motives and basic traits. Below we attempt to delineate these two constellations of variables, which we titled Alpha and Gamma, and then integrate them into a two-factor model. We then go on to review some empirical work consistent with the model.

Over 30 years ago, Wiggins (1964) and Block (1965) applied the labels “Alpha” and “Gamma” to distinguish two social desirability-laden dimensions resembling those that we address here. Unfortunately, this important distinction was not pursued, largely, it appears, because commentators could not agree whether Alpha represented distortion (Edwards, 1957; Messick & Jackson, 1958) or personality content (Block, 1965). Believing that both characterizations apply, we see the noncommittal nature of the labels as a virtue. And we find the labels apt for at least two other reasons. First, they were proposed by two influential contributors who labored at the intersection between personality and
self-evaluative processes. Second, the use of connotation-free Greek names requires, as yet, no commitment to interpretation.

With this two-factor template to guide our literature review, we pinpointed Alpha and Gamma themes in research on personality traits, motives, and defensive mechanisms. The breadth of their application may be seen in the four sections of Figure 2, that is, Response Styles, Self-Evaluation and Self-Favoring, Values and Motives, and Self-Perceived Personality.

**Social Desirability Response Styles**

As we noted above, early work on socially desirable responding emphasized one factor (Edwards, 1957; Jackson & Messick, 1958). It was Wiggins (1964) who factor-analyzed all the available measures of socially desirable responding and named the two distinct forms “Alpha” and “Gamma.” Damarin and Messick (1965) interpreted them as autistic bias and propagandistic bias. Some time later, Paulhus (1984) factored a wider array of measures yielding essentially the same factors. Unlike earlier noncommittal application of Greek letters, Paulhus concluded that the two dimensions should be interpreted as self-deception and impression management.

In a later refinement, Paulhus and Reid (1991) concluded that the self-deception items could be partitioned into two distinct factors: This research led to the development of two relatively orthogonal subscales, Self-Deceptive Enhancement and Self-Deceptive Denial (Paulhus, 1988, 1998b). Although both tendencies show demonstrable distortions in self-perception, neither is responsive to audience effects.

Interestingly, the Denial scale, but not the Enhancement scale, is positively and highly correlated with the Impression Management scale. The fact that denial responses (“I never enjoy my bowel movements”) and impression management responses (“I always pick up my litter”) tend to co-occur suggests that the same individuals tend to exhibit both conscious and unconscious versions of exaggerated moralism and social conventionality—at least under anonymous conditions.

In sum, the current literature indicates two forms of self-deceptive socially desirable responding, interpretable as self-enhancement and denial. That is, people can unconsciously distort the favorability of

---

8 This work was later published in a more accessible medium (Messick, 1991).
<table>
<thead>
<tr>
<th><strong>Self-Deceptive Styles: Egoistic &amp; Moralistic Bias</strong></th>
<th>1037</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response Styles</strong></td>
<td><strong>ALPHA</strong></td>
</tr>
<tr>
<td>Damarin/Messick (1965)</td>
<td>autistic bias</td>
</tr>
<tr>
<td>Paulhus (1984)</td>
<td>self-deception</td>
</tr>
<tr>
<td>Edwards (1970)</td>
<td>social desirability</td>
</tr>
<tr>
<td>Roth/Snyder/Pace (1986)</td>
<td>attributive</td>
</tr>
<tr>
<td>Paulhus/Reid (1991)</td>
<td>self-deceptive enhancement</td>
</tr>
<tr>
<td><strong>Self-Evaluation and Self-Favoring</strong></td>
<td><strong>Messick (1960)</strong></td>
</tr>
<tr>
<td></td>
<td>narcissistic exhibitionism</td>
</tr>
<tr>
<td>Rosenberg/Sedlack (1972)</td>
<td>intellectually good</td>
</tr>
<tr>
<td>Jones and Pittman (1982)</td>
<td>self-promoter</td>
</tr>
<tr>
<td>Higgins (1987)</td>
<td>ideal-self</td>
</tr>
<tr>
<td>Millon (1990)</td>
<td>narcissistic personality disorder</td>
</tr>
<tr>
<td>Raskin/Novacek/Hogan (1991b)</td>
<td>grandiose self-enhancement</td>
</tr>
<tr>
<td>John/Robins (1994)</td>
<td>narcissism</td>
</tr>
<tr>
<td>Paulhus/John (1994)</td>
<td>superheroes</td>
</tr>
<tr>
<td>Tellegen/Waller (1987)</td>
<td>positive valence</td>
</tr>
<tr>
<td><strong>Values and Motives</strong></td>
<td><strong>Freud (1894/1953)</strong></td>
</tr>
<tr>
<td></td>
<td>superiority-striving</td>
</tr>
<tr>
<td></td>
<td>need for superiority</td>
</tr>
<tr>
<td></td>
<td>approach success</td>
</tr>
<tr>
<td></td>
<td>agency</td>
</tr>
<tr>
<td></td>
<td>individualism</td>
</tr>
<tr>
<td></td>
<td>dominance</td>
</tr>
<tr>
<td></td>
<td>achievement</td>
</tr>
<tr>
<td></td>
<td>getting ahead</td>
</tr>
<tr>
<td></td>
<td>competence</td>
</tr>
<tr>
<td></td>
<td>independence</td>
</tr>
<tr>
<td></td>
<td>self-enhancement</td>
</tr>
<tr>
<td></td>
<td>competence</td>
</tr>
<tr>
<td></td>
<td>need for power</td>
</tr>
<tr>
<td><strong>Self-Perceived Personality</strong></td>
<td><strong>Block (1965)</strong></td>
</tr>
<tr>
<td></td>
<td>ach’t via independence</td>
</tr>
<tr>
<td></td>
<td>dominance</td>
</tr>
<tr>
<td></td>
<td>social potency</td>
</tr>
<tr>
<td></td>
<td>dominance</td>
</tr>
<tr>
<td></td>
<td>perceived capability</td>
</tr>
<tr>
<td></td>
<td>(lack of) distress</td>
</tr>
<tr>
<td></td>
<td>personal growth</td>
</tr>
<tr>
<td></td>
<td>socialization</td>
</tr>
</tbody>
</table>

**Figure 2**

Manifestations of Alpha and Gamma Constellations at Four Levels of Personality
questionnaire responses in two ways: by exaggerating their talents and by minimizing their sins (for a review, see Paulhus, in press).

**Self-Evaluation and Self-Favoring**

One-dimensional models of self-evaluation have long dominated the self-perception literature (e.g., Anderson, 1968; Edwards, 1957; Osgood, Suci, & Tannenbaum, 1957; Peabody, 1984; Saucier, 1994; Taylor & Brown, 1988). Overlooked were a pair of factors that recurred consistently in work exploring qualitatively different forms of evaluation. Messick (1960), for example, found multiple points of view in how subjects rated the desirability of traits: The two largest factors were labeled “narcissistic self-promotion” and “Protestant ethic.” Using multidimensional scaling, Rosenberg and Sedlak (1972) found that two evaluative dimensions resulted from scaling the perceived similarity of trait adjectives: One was interpreted as “intellectually good” (e.g., intelligent, determined) and the other as “socially good” (e.g., reliable, helpful). Subsequent research documented the same two primary factors (e.g., Kinder & Sears, 1985; Wojciszke, 1994).

In the domain of self-reported personality, a parallel pair of evaluative factors have been isolated (Benet & Waller, 1995; Tellegen & Waller, 1987). The first, labeled Positive Valence, contrasts narcissism-flavored items such as “outstanding” and “superior” with “normal” and “average.” The second factor, labeled Negative Valence, emphasizes extreme and socially deviant characteristics such as “cruel” and “deserves to be hated.” Both represent exaggerated forms of self-evaluation different in tone from the common view of high self-esteem. And both resemble the factors of self-favoring bias emerging from our own factor analyses of self-criterion residual (SCR) scores, namely, an egoism regarding one’s status and prominence, on the one hand, and a moralism regarding one’s virtue as a dutiful and agreeable person, on the other (Paulhus & John, 1994).

To summarize: Contrary to the earlier view that evaluation is unidimensional, the work reviewed here documents the repeated emergence of the same two dimensions in research on self-evaluations and self-favoring. One dimension reflects individual differences in the narcissistic, egoistic tendency to claim exceptional qualities, that is, superior status and competence. The second factor reflects an exaggerated tendency
to see oneself as a good person, that is, dutifully adhering to social norms and avoiding antisocial behavior.

**Values and Motives**

Disputes about the number of important human values have a long history. Yet two, above all, are never overlooked on lists of values. They are agency and communion (Bakan, 1966), or variants thereof. Agency is the positive value placed on individuality, personal striving, growth, and achievement. Communion is the positive value placed on relationships, intimacy, and benefiting others, even the society as a whole. Similar dimensions appear as the latent structure underlying the most comprehensive, cross-cultural analyses of values (Bond, 1988; Lonner, 1980; Schwartz, 1992). In the argot of everyday life, the values have been felicitously labeled as “getting ahead” and “getting along” (Hogan, 1983).

These two broad cultural values are assumed to promote the socialization of two sets of value-consistent traits. Different histories of socialization, then, play a major role in the dramatic individual differences observed in these traits. As a result, researchers in the personological tradition have consistently emphasized in their lists of needs some version of power and affiliation (McAdams, 1985; Murray, 1938; Winter, 1996). The dual-motive framework has also been highlighted in the extensive empirical and theoretical work by Wiggins and his colleagues: There, agency and communion are held to be the primary motives giving rise to personality traits (Wiggins, 1991; Wiggins & Trapnell, 1996).

In short, when limited to two values, conceptual and empirical analyses usually yield constructs akin to agency and communion. Individual differences in these values then induce observed differences in the motives of need for power and need for approval.

**Self-Perceptions of Personality**

Under most circumstances, comprehensive analyses of self-report personality items exhibit five factors (e.g., McCrae & Costa, 1987; John, 1990).

---

9. Alternatively, the two dimensions may derive from distinctive adaptive strategies. In this view, evolution has endowed humans with two powerful, but independent, motives that enhance reproductive success (Buss, 1997; Gangestad & Simpson, 1990). Genetic variation on these motives may explain the prominence of agentic and communal traits.
Over the years, however, there have been repeated indications of small but systematic intercorrelations among the five factors. This pattern led Digman (1997) to reanalyze several diverse data sets, by factor-analyzing the intercorrelations among Big Five measures. He concluded that the five factors can be organized into two themes or superfactors, namely, Personal Growth (Extraversion and Openness) and Socialization (Agreeableness, Conscientiousness, and Emotional Stability). Drawing on Digman’s data and other literature, Wiggins and Trapnell (1996) also argued for the primacy of two themes underlying the Big Five dimensions. According to their analysis, the facets of each Big Five dimension can be linked to either agency or communion.

Finally, it is not uncommon to find (sometimes substantial) intercorrelations among the Big Five in select samples or under select conditions. The possibility that a high demand for desirable responding might be responsible for these intercorrelations is supported by experimental evidence showing that high intercorrelations can be induced by instructions to respond in a socially desirable fashion (Paulhus, Bruce, & Trapnell, 1995). But this lack of orthogonality sometimes appears under standard “straight-take” conditions. Nonetheless, the operation of a social desirability response set has been cited as one possible cause for the sometimes exaggerated intercorrelations among Big Five self-reports (e.g., Digman, 1997; McCrae & Costa, 1989, p. 435). To summarize: Under certain conditions, nonzero intercorrelations are observed among measures of the Big Five traits. This systematic and replicable convergence suggests the intrusion of two evaluative superfactors.

### Integrating Each Constellation

According to Anna Freud, where there is motive, there is distortion and defense (1936). Perhaps it is coincidence that similar pairs of dimensions are implicated whether we look at response styles, personality, motives, or self-evaluation. We think not. Instead, we propose that Alpha and Gamma are two constellations of values, motives, and biases that influence self-perceived personality. Our hypothesis is that they interrelate in

10. Although his factors bear some resemblance to ours, Digman (1997) used the labels Alpha and Beta to name the factors that we would label Gamma and Alpha. He assigned the names arbitrarily (personal communication) whereas we followed the historical precedents set by Block (1965) and Wiggins (1964).
the sequence depicted in Figure 3: Two fundamental human values, agency and communion, give rise to two motives, need for power and need for approval. The latter motives dispose some individuals to an egoistic tendency to exaggerate agentic qualities and a moralistic tendency to exaggerate communal qualities, respectively. Individual differences in these two self-favoring tendencies, in turn, influence self-perceptions of personality in a systematic fashion. The egoistic and moralistic biases are both self-deceptive styles in that they operate unconsciously to preserve and amplify a positive self-image. In this model, then, values lead inexorably to biases.¹¹

Factor Alpha

We hold that the list of concepts under the Alpha factor in Figure 2 are all connected with a self-favoring bias that we think is best captured by the term “egoistic bias.” This mechanism entails an exaggerated self-worth with regard to social and intellectual competence. It is played out in a self-favorability bias influencing how people perceive themselves on personality dimensions related to competence and social status (see Figure 1). Empirically, the concept is closely related to “normal narcissism” as operationalized by the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1981). For example, John and Robins (1994) found that the NPI (and three other measures of overt narcissism) predicted the degree to which subjects overestimated how well they performed on a

<table>
<thead>
<tr>
<th>Value</th>
<th>Motive</th>
<th>Self-Deceptive Mechanisms</th>
<th>Self-Favoring Bias on</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALPHA</td>
<td>agency</td>
<td>need for Power</td>
<td>egoistic bias</td>
</tr>
<tr>
<td>GAMMA</td>
<td>communion</td>
<td>need for Approval</td>
<td>moralistic bias</td>
</tr>
</tbody>
</table>

Figure 3

The Constellations of Alpha and Gamma

¹¹. Of course, there are other plausible sequences. We placed biases at the end of the sequence primarily because they are the focus of attention here.
group task, as indexed by three different criteria (see Robins & John, 1997a, 1997b).

Such measures of normal narcissism overlap conceptually and empirically with a measure designed specifically to tap egoistic bias, namely, Paulhus and Reid’s (1991) Self-Deceptive Enhancement scale. Typically, the two measures intercorrelate in the .50 range (McHoskey, Worzel, & Szyarto, 1998; Paulhus, 1988). More important, Paulhus (1998a) found a complex pattern of observer-rating correlates of the NPI that were remarkably similar to those of the Self-Deceptive Enhancement scale.

**Factor Gamma**

We propose that the list of concepts listed under Gamma in Figure 2 is associated with a bias best described as moralistic. This second self-deceptive style serves a self-protective function, best indexed by the Marlowe-Crowne or Impression Management scales. This style entails an avoidance of disapproval by conforming to social norms. Weinberger’s notion of “restraint,” one of two primary dimensions underlying psychopathology, is closely related. In fact, Weinberger’s original operationalization of this dimension was the Marlowe-Crowne scale (Weinberger et al., 1979). Later, he developed a broader measure (Restraint) with items better capturing the personality aspects, rather than the bias aspects, of this dimension (see Weinberger, this volume).

The nature of the Gamma factor may be further clarified by an examination of the self-evaluative dimension with which it is strongly (and negatively) associated, namely, Negative Valence (Tellegen & Waller, 1987). The items on this measure consist primarily of socially deviant behaviors (cruel, violent, immoral). The moralistic disapproval that humans feel toward such characteristics is essential to society because they threaten affiliation and harmony. In other words, such behaviors threaten society itself by undermining the communal forces that keep it unified.

We suggest that excessive motivation in this domain, however, may lead to (1) exaggerated “saint-like” self-conceptions (Paulhus & John, 1994), and (2) attempts to ensure a Gamma-like image in the eyes of others (Millham & Jacobson, 1978). This tendency should be manifested in self-perceptions on personality dimensions such as Agreeableness and Dutifulness.
Some Empirical Evidence for the Alpha-Gamma Model

Recently, we put the Alpha-Gamma Model to the test in two ways. One approach analyzed the relation of self-favorability factors to other possible indicators of Alpha and Gamma. The second approach examined the structure of the Big Five under high and low demand for self-presentation.

Factor Analytic Studies

To further test the two-dimensional structure assumed to underlie the elements in Figure 2, we attempted (1) to replicate the factor structure of the self-criterion residuals shown in Figure 1 and (2) to relate the resulting factor scores to direct measures of as many as possible of the variables related to Alpha and Gamma. Here we can provide only a review of a series of such factor analyses.

The data sets were two large samples of both self- and peer ratings on the Big Five Inventory (John, Donahue, & Kentle, 1991), a well-validated measure of the Big Five trait domains. First, we confirmed that the self-rating data and the peer-rating data showed the usual five factors. Then, for each factor, we indexed self-favorability as the residual self-rating after all peer rating variance has been removed. This procedure yielded five self-favorability scores that we intercorrelated. A factor analysis of this matrix yielded two self-favorability factors, one marked by Extraversion and Openness and the other marked by Agreeableness and Conscientiousness. In short, this result replicated the factors labeled “Saints” and “Superheroes” by Paulhus and John (1994). These findings suggest that the structure of self-favorability measures (2-D) is noticeably smaller than the structure of the Big Five (5-D) appearing in both the self-rating and peer-rating domains.

We computed factor scores for those two factors and examined their associations with a number of the independent indicators of Alpha and Gamma listed in Figure 2. Regression analyses showed that the two self-favorability factors could be predicted from the Alpha and Gamma indicators with substantial multiple Rs. The simplest way to summarize these associations is to project them onto the two-space defined by the self-favoring bias factors. As Figure 4 shows, the external correlates marking Factor 1 included the Narcissistic Personality Inventory, Self-Deceptive Enhancement, and Positive Valence. External correlates
marking Factor 2 included Self-Deceptive Denial and Negative Valence (reverse-scored), and (to a lesser extent) the Marlowe-Crowne scale. A comparison of Figures 1 and 4 shows that biases in self-perceived personality are differentially related to independent measures of personality and self-evaluation. More generally, the findings begin to flesh out the coherence of the Alpha and Gamma patterns we have proposed in Figure 2.

In short, the two-factor solution for the five self-favoring bias scores yielded the predicted Alpha and Gamma structure; this solution contrasted dramatically with the five-factor solution required for either the self-ratings or peer-ratings factored separately.

Faking Studies

Another series of studies examined the structure of self-favorability scores under different levels of demand for self-presentation (Paulhus & Notareschi, 1993). Subjects were asked to respond to standard personality inventories under varying instructions to create a socially desirable impression (see Paulhus et al., 1995). Factor analyses showed a reduction
in dimensionality from five to two factors under high demand. Moreover, the factors appearing under high demand closely resembled the two factors of self-favorability found by Paulhus and John (1994), namely, Alpha and Gamma. That is, Agreeableness and Conscientiousness became much more highly correlated with each other (indicating Gamma) whereas Extraversion, Neuroticism, and Openness became more highly intercorrelated (indicating Alpha).

**Summary and Implications**

We have argued that self-reports on normal personality traits are intertwined with two self-deceptive styles, that is, egoistic bias and moralistic bias. We hypothesize that these styles, in turn, derive from two motives—the need for power and status and the need for approval—which result from the fundamental human values of agency and communion. Together, these two constellations of associated values, motives and biases have been labeled Alpha and Gamma.

**The Scope of Alpha and Gamma**

*Egoistic Bias.* This self-deceptive style entails an exaggerated self-worth with regard to social and intellectual status. This self-perception has a “superhero” quality (Paulhus & John, 1994). Our research suggests that individual differences are best operationalized by normal narcissism scales such as the NPI (Raskin & Hall, 1981) or the Self-Deceptive Enhancement scale (Paulhus, 1991). Elsewhere, we have shown that high scorers on these measures have short-term advantages of self-confidence and positive first impressions (Paulhus, 1998a). In the long term, however, both their interpersonal appeal and their mental health are evaluated negatively by others (Robins & John, 1997a; Paulhus, 1998a).

The effects of an egoistic bias are played out in exaggerated self-perceptions on personality dimensions of Extraversion and Openness, as well as on ability dimensions of intelligence and creativity. These attributes all involve agentic qualities, that is, prominence and status. Put another way, such individuals seek to stand out from the crowd.

Future research may show that this pattern varies across specific facets of these variables (see McCrae & Costa, 1995). For example, egoistic bias may operate on (1) the Extraversion facets of assertiveness, activity,
and excitement seeking, but not on sociability; (2) the Openness facets of fantasy, ideas, and actions, but not on feelings; (3) the Neuroticism facets of depression and self-consciousness (both reverse scored); and (4) those Conscientiousness facets that involve competence and achievement-striving. These predictions are consistent with recent speculation by Wiggins and Trapnell (1996) that facets related to agency and communion facets may be found within Conscientiousness, Emotional Stability, and Openness, the last two factors of the Big Five. The first two factors of the Big Five, of course, map directly onto agency and communion, respectively.

Moralistic Bias. This self-deceptive style entails an exaggerated self-positivity with regard to being a “nice person” and a “good citizen”—in short, good character. With sufficient exaggeration, such self-perceptions reveal a sanctimonious “saint-like” self-image. These tendencies are best measured by the Self-Deceptive Denial, Impression Management, or Marlowe-Crowne scales. Although some high scorers on these measures strive for and indeed even achieve many desirable communal attributes, they can also exaggerate and lie, if necessary, to ensure that others recognize this saintly impression (Millham & Jacobson, 1978).

Moralistic bias is played out in inflated scores on traits related to Agreeableness (cooperativeness, warmth) and self-control aspects of Conscientiousness (e.g., responsibility, dutifulness, order). These traits all involve avoidance of disapproval by adhering to social norms and persevering at assigned tasks.

Some people strive to be “saints”: They believe they can control earthly impulses to a greater degree than their morally weaker peers. Others strive to be “superheroes”: They believe they can accomplish great feats that are unattainable by their contemporaries (Damarin & Messick, 1965). And we do not deny that both types fulfill their ideals to a certain degree.

Broader Implications. Other writers have argued for the fundamental importance of Alpha and Gamma in terms of their underlying values, agency and communion (e.g., McAdams, 1985; Wiggins, 1991). Not only are both forces adaptive, but a tension between the two forces allows for an optimal balance of goal-striving activities within the individual and within societies. The bias mechanisms can contribute to single-minded
striving observed in some individuals: Those who see themselves as superheroes or saints can both be influential.

At a macro level, Alpha and Gamma map onto several larger dynamic tensions in contemporary society. One correspondence is with the balance of political forces typical of modern, flourishing societies, namely, capitalism and socialism (Kerlinger, 1984). That is, growth in economic systems is said to be driven by the “engine” of capitalism whereas the “human face,” that is, the concern with those suffering from the system, is provided by socialist activism. Another relevant contrast, especially between Western and Eastern cultures, is often couched in the parallel labels of individualism-collectivism (Triandis, 1978) or independence-interdependence (Markus & Kitayama, 1991). Given such a clear connection between fundamental biases and political attitudes, it is not surprising, then, that the biasing effects of political orientation are so powerful (Paulhus, 1996; Tetlock, 1994).

Finally, it may not be a coincidence that Alpha and Gamma are consistent with traditional sex-role distinctions. After all, the traditional male role emphasizes individual achievement outside the home whereas the female role emphasizes maintaining harmony in home, family, and relationships. Not surprisingly, then, the clearest sex differences in traits correspond to these value differences: Males score higher on agency-related values and traits, and women score higher on communal-related values and traits (Helgeson, 1997; Schwartz, 1992; Wiggins, 1991). Similarly, men score higher on egoistic biases (Aube & Koestner, 1994; Beyer & Bowden, 1997) and women score higher on Gamma biases (Paulhus, 1988, 1998b).

12. Rhodewalt and Morf (1995) demonstrated the low cognitive complexity of those with an egoistic bias (narcissists). We predict that individuals with a moralistic bias will similarly exhibit low cognitive complexity.
13. Given the contrast in socialization, Gilligan (1982) has gone so far as to argue that men and women’s moral development should be evaluated on different criteria, namely, achievement-orientation and relationship-orientation, respectively. Furthermore, well-being in men and women appears to depend more on agency and communion, respectively (Helgeson, 1994).
14. In fact, Bem’s (1974) influential Masculinity and Femininity scales have been shown to correspond almost directly with measures of agency and communion (Wiggins & Holzmuller, 1981).
Mechanisms for the Biases

Levels of Consciousness. The present model departs substantially from that espoused earlier by one of us (Paulhus, 1986). The earlier model claimed that the critical difference between the two major forms of self-report bias was the level of conscious awareness, with Gamma-like measures implying more conscious forms of self-serving distortion and impression management. We now interpret Alpha and Gamma differences as referring to different content, that is, agentic and communal aspects of personality. Moreover, Alpha and Gamma constellations include both conscious (impression management) and unconscious (self-deceptive) forms of bias.

At least two pieces of research justify this reinterpretation. First, correlations between self-reported personality and Alpha and Gamma bias measures are evident even under anonymous conditions where conscious efforts at impression management are implausible (Paulhus & John, 1994). Second, our research has shown that both Alpha bias and Gamma bias can be manipulated with appropriate instructions (act agentic vs. act communal) (Paulhus & Notareschi, 1993; Paulhus, in press). In retrospect, we suspect that, in past studies, typical faking instructions tended to have more effect on Gamma measures because the instructions typically emphasized communal rather than agentic qualities (e.g., “give the socially desirable response”), thus invoking the “morally good” attributes of the good citizen and worker.

Given that the self-ratings in our own SCR data (see Figure 1) were relatively anonymous, we suggest that both factors involve self-deceptive (i.e., nonconscious) processes. When combined with the faking studies, the results suggest that both Alpha and Gamma characteristics can be consciously and unconsciously inflated. In general, the unconscious bias should push in the same direction as the trait: One reason is that self-perceptions are often exaggerations of a kernel of truth. For example, highly agreeable people strive to be saint-like but, as the peer ratings suggest, they do not succeed quite to the extent that they would wish. Thus individuals are also more likely to show self-perception biases in the same direction as their traits. However, rather than five independent self-perception biases, one for each of the Big Five, we found evidence for only two underlying biases.

The conscious versions of Alpha and Gamma bias (Alpha and Gamma management) are likely to be more flexible and tailored to suit the
audience (Paulhus, in press). Conscious promotion of one’s assets is usually labeled “bragging.” Conscious defense of one’s public image or “good name” is commonly labeled “impression management” (Jones & Pittman, 1982). These conscious versions are less relevant to this article than are the self-deceptive components, but we plan to pursue this form in future research. They are of particular relevance in addressing recent claims that conscious distortion is of no concern in job applications (Ones, Viswesvaran, & Reiss, 1996).

*Competence Domain versus Moral Domain.* The critical difference between Alpha and Gamma bias measures now appears to correspond to the different values and motives of the Alpha and Gamma constellations. Different biases derive from valuing different goals: To the extent that one values status and prominence, then an egoistic bias is more likely (although certainly not necessary). And to the extent that one values cooperation and dutifulness, one may overlook one’s own deficiencies in that regard.

It can be argued that both biases are defensive styles (see Raskin et al., 1991a). Following classic explanations of narcissism originating with Freud, the observed self-promotion derives from deep-seated insecurity. If a child’s parents provided attention and affection only upon demonstrations of competence, it is easy to understand how the consequent adult character would show an excessive need to self-promote. Similarly, if withdrawal of affection followed any deviation from parental rules, then excessive concern with being “naughty” could dominate one’s adult character (Rank, 1945).

*Attribution Versus Denial.* An alternative to the content-driven explanation for the distinction between Alpha and Gamma bias is the distinction between claiming positive attributes and denying negative ones (Sackeim, 1983; Roth, Snyder, & Pace, 1986). Both are self-favoring strategies, but the attributional form focuses on claiming rare positive qualities whereas the denial form emphasizes never having deviated from common good behaviors. Presumably individual differences in these dimensions would correspond to individual differences in (1) sensitivity to reward and (2) sensitivity to punishment. Sackeim (1983) made this same argument in developing his claim that self-deception can be either offensive or defensive in function.
Differences in content might then follow from differences in function. After all, attribution is only self-favoring if one claims to have rare, exceptional qualities. Such individuals will exaggerate qualities that will raise their social status and their social prominence, that is, power, intellect, and extraversion. With a Gamma bias, being "good" is construed as being normal or appropriate, that is, not standing out from the crowd. Hence, the self-favoring process will involve the denial of socially unacceptable aspects of the self, for example, antisocial behavior such as cruelty and dishonesty and Id-related impulses such as aggression toward parents and authorities and sexual indiscretions. This notion is consistent with the operation of a "feared self" (Markus & Nurius, 1986; Sullivan, 1953) as well as with traditional psychoanalytic ideas (Cramer, 1995).

The Structure of Self-Perceived Personality

When correlations are observed among the Big Five, they are systematic. For example, Agreeableness and Conscientiousness will often correlate positively (Digman, 1997; Eysenck, 1993) and Extraversion and Openness often show positive correlations (McCrae & Costa, 1989; Goldberg, 1993).

In the faking studies reported above, such correlations were induced by a high demand for self-presentation. Although instructions in those studies aimed to promote a global self-positivity (i.e., "fake good"), we suspect that Alpha and Gamma biases can be activated independently. Gamma-bias activation might occur in public or family contexts where avoiding shame and loss of affection are the primary concerns. Under such circumstances, people will be cooperative with peers and obey authorities. Alpha-bias activation, in contrast, may occur in competitive circumstances where potential glory or sexual partners are the primary concern. Here people will exaggerate their boldness, risk-taking, and creativity.

The Self-Criterion Residual (SCR) Method

Development of the Alpha-Gamma theory was greatly facilitated by our new method for evaluating the structure of self-favoring biases. The method was modeled on John and Robins’s (1994) residual measure of self-enhancement. We extended their single-measure analysis to a comprehensive collection of many such measures and examined their
structure. Our long-range plan is to explore the utility of the SCR technique in resolving the conundrum of content versus style.

The SCR method certainly has more face validity than traditional self-report measures of social desirability response styles because it indexes directly the degree to which the self-perception deviates from scores on an independently assessed criterion, such as reports by well-informed observers or other relevant criteria. As a result, it provides a model for the self-deceptive process underlying defense mechanisms. According to Sackeim and Gur (1978), the self-deceptive process requires two conflicting representations of the same information. In SCR, the two representations of reality are that of the self and that of a (more) objective criterion such as clinical ratings or behavior recorded on-line by objective observers (Gosling, John, Craik, & Robins, 1998).

Consider the recent study by Robins and John (1997b), however. When narcissists were shown videotapes of their actual performance, their inflated self-evaluations actually increased, rather than decreased. Such research suggests that narcissists are not merely overlooking objective criterion information but actively distorting it. Further research of this sort should be fruitful in uncovering the mechanisms behind such motivated distortions.

**Relation to Weinberger Model**

There is some similarity between our Alpha-Gamma Model and the two-factor theory of defense developed by Weinberger (1990, this volume). Like us, Weinberger argued that there are two fundamental dimensions in personality that bear on defensive processes; for him, they are trait differences in distress and self-restraint. To analyze defensive style, he partitions distress into two levels and restraint into three, yielding six “typological” combinations. Although one particular combination, low distress and high restraint (repressors), receives special consideration, the two dimensions themselves have implications for defenses as well (Weinberger, this volume).

It is certainly possible to draw a parallel between our Alpha and Gamma and Weinberger’s dimensions of (low) Distress and Restraint. Our Gamma dimension is clearly similar to Restraint in describing high scorers as oversocialized. Moreover, low scorers on his Distress dimension bear a similarity to high scorers on our Alpha dimension (in fact, distress measures correlate highly negatively with Alpha, consistent with
narcissists’ low levels of self-reported anxiety and distress). The clear
difference is that we emphasize the self-deceptive aspect of each dimen-
sion. So how do Weinberger’s repressors fall out in our system? They are
likely to be high on both Alpha and Gamma: that is, they are both egoistic
and moralistic. Being self-deceptive in two ways, they are doubly prone
to the advantages and disadvantages of self-deception (see Lockard &
Paulhus, 1988).

CONCLUSIONS

We hope that we have convinced the reader that defensive biases intrude
into self-perceptions of personality and ability. The reader might even
find some credibility in our more ambitious proposition that values and
motive serve as antecedents in this sequence. To recap: Alpha is our label
for the sequence starting with individual differences in the tendency to
value agency, that is, social status and prominence. These differences
then induce differences in motivation to excel; differences in this motive
may then alter self-perceptions of intelligence, extraversion, and open-
ness. Similarly, Gamma is our label for the sequence beginning with
individual differences in valuing communion, that is, membership in
one’s social group. Value differences then induce differences in need for
approval, which in turn may induce a moralistic bias. This bias, in turn,
distorts self-perceptions of conscientiousness and agreeableness.

We would argue that, empirically, the case is strongest for the connec-
tion between corresponding traits and self-deceptive biases. Badly lack-
ing, however, is research connecting measures of our egoistic and
moralistic biases with measures of traditional defense mechanisms (e.g.,
Davidson & McGregor, 1996; Gleser & Ihilevich, 1969). With regard to
the antecedent role of values and motives in the Alpha or Gamma
sequences, the empirical case is yet to be made. Research that induces
or, at least, triggers agency and communion motives is necessary to make
that case. Among the possibilities is research involving the activation of
one motive or the other by subtle means such as trait priming (Bargh &
Pratto, 1986) or ought-self versus ideal-self priming (Higgins, Roney,
Crowe, & Hymes, 1994).

In the end, the prudent reader may hesitate to grant the full sweep
of our arguments. But an acceptance of our central claim—that there
are egoistic and moralistic forms of self-favoring bias—goes a long
way toward a resolution of several contentious issues in personality
psychology. These include understanding: (1) the conceptual basis of bias in self-reports, (2) how to operationalize these biases, (3) the distinction between promoting and defending one’s self-image, and (4) why the (orthogonally extracted) Big Five factors often intercorrelate. If our lofty claims seem to depart from reality, however, egoistic biases may be at fault.

REFERENCES


Self-Deceptive Styles: Egoistic & Moralistic Bias


