



Contrasting perspectives on personality problems: descriptions from the self and others

Allan Clifton, Eric Turkheimer^{*}, Thomas F. Oltmanns

Department of Psychology, University of Virginia, P.O. Box 400400, Charlottesville, VA 22904-4400, USA

Received 12 August 2002; received in revised form 27 March 2003; accepted 8 June 2003

Abstract

Self-reported personality disorder (PD) traits have low to moderate correlations with informant and peer-reports of the same traits. The current study investigates the source of disagreement in correspondence and proposes that in some cases it is due to systematic differences in perceptions of traits by self and others. Military recruits ($N = 2013$) were administered self-report measures and an inventory in which they nominated peers who displayed DSM-IV PD traits. Correspondence between self- and peer-report for the same diagnostic categories was low (R s ranging from 0.11 to 0.22). Semipartial correlations were used to find the 10 items which best predicted each diagnostic category over and above the corresponding scale. Adding these supplemental items to the original regression significantly increased the amount of variance explained (R s ranged from 0.15 to 0.35) in a cross-validation sample. Although peer- and self-report differed in content, the relationship between self and peer perceptions appears to be systematic and meaningful (e.g., people whose peers describe them as paranoid describe themselves as angry and hostile).

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Self-report; Peer-report; Personality disorders

1. Introduction

Personality disorders (PDs) are most often diagnosed on the basis of self-report, obtained through written inventories or clinical interview, which may produce a limited view of PD traits. Self-report involves only one source of information, limiting its psychometric utility. Further, the nature of PDs inherently involves the way in which one's maladaptive personality affects others (Westen, 1997). Persons with PDs may have an especially difficult time viewing themselves

^{*} Corresponding author. Tel.: +1-434-982-4732; fax: +1-434-982-4766.
E-mail address: turkheimer@virginia.edu (E. Turkheimer).

realistically, or accurately perceiving the effects their behavior has on others (Oltmanns, Turkheimer, & Strauss, 1998). The use of informants provides additional sources of information, as well as an alternative view of the interpersonal aspects of personality pathology. However, peer- and self-report may produce different pictures that are hard to reconcile.

1.1. Methods of peer assessment of personality

Peer perceptions of pathological personality traits are usually obtained from one or more selected informants (McCrae & Costa, 1987; Zimmerman, Pfohl, Stangl, & Corenthal, 1986). The participant being rated is asked to identify one or more persons knowledgeable about him or her, such as a spouse, relative, or close friend. The informant is then asked to describe the personality of the participant via questionnaires, structured interviews, or other means. This informant information is compared with self-report personality data obtained from the participant, and the congruence between the two sources of information is determined.

Information can also be obtained from larger unselected peer groups that interact with the individual on a regular basis, such as co-workers or classmates (Mount, Barrick, & Strauss, 1994; Oltmanns et al., 1998). Multiple informants have consistently shown a high degree of inter-rater reliability (Funder, 1991; Harris & Schaubroeck, 1988; Kenrick & Funder, 1988; McCrae & Costa, 1987). The interjudge correlations of acquaintances, even those who know the target from different contexts, are significantly higher than those of strangers (Funder & Colvin, 1988; Funder, Kolar, & Blackman, 1995), suggesting that knowledge of the target, rather than informed guesses or stereotypes about how personality hangs together, are necessary for reliable personality judgments.

Gathering information from a large group can be performed by either nomination or rating (or by rank ordering, which is less frequently done). In nomination, peers are asked to identify a set number of members of the group who best demonstrate a particular criterion. In ratings, peers are asked to rate each member of the group as to how much he or she demonstrates the criterion. Nominations are more practical for large groups, and demonstrate higher reliability, predictive validity (Kane & Lawler, 1978), and discriminant validity (Schwarzwald, Koslowsky, & Mager-Bibi, 1999) than ratings. However, nominations tend to provide more information about the extreme members of groups, and less about those in the middle, which may artificially increase estimates of validity (Kane & Lawler, 1978). Peer ratings, in which each member is separately rated on each trait, are useful in obtaining more general information about all individuals (Maassen, Goossens, & Bokhorst, 1998), and may have higher test-retest reliability (Asher & Hymel, 1981; Putallaz & Gottman, 1981). The present study utilizes a hybrid nomination-rating scale, in which peers nominate an unlimited number of those who best fit the criteria, but also provide a rating of those they nominate.

1.2. Self-other correlations in normal personality

Numerous studies have investigated consensus between self-reported normal personality traits and those reported by peers or informants. Several studies utilizing the NEO-PI have found low to moderate self-informant correlations, with *R* values ranging from 0.1 to 0.6 (Funder et al., 1995; McCrae & Costa, 1987). Similarly, Fuhrman and Funder (1995) found moderate to strong self-

other correlations for the California Personality Inventory (mean r ranging from 0.52 to 0.68). Studies using the Q-sort technique of personality assessment have demonstrated overall self–peer correlations around 0.3, although this varies widely across individual items (Funder, 1991; Funder & Colvin, 1988).

The correlations among informants' ratings are generally higher than the correlations between informants and self-ratings (Kenny, 1994). Informants' personality ratings also predict subsequent behavior at least as well as self-ratings (Hurley, 1988; Levesque & Kenny, 1993; Mount et al., 1994). Kolar, Funder, and Colvin (1996) found that aggregate ratings by multiple informants actually correlated higher with observed behavior than did self-ratings. Data from our own study corroborate this finding. We have used peer- and self-report data that were collected at the end of basic military training to predict early discharge from the Air Force. Peer nominations for antisocial features were significantly correlated with the tendency for recruits to be separated from the military within three years of enlistment (Fiedler, Oltmanns, & Turkheimer, in press).

1.3. Self-other correlations in personality disorders

PD studies of selected informants have found self-other agreement similar to that seen in studies of normal personality. A recent review of informant reports of PDs (Klonsky, Oltmanns, & Turkheimer, 2002) found a median self-informant correlation of 0.36 in studies of DSM PDs. Studies of non-DSM personality pathology, such as SNAP trait scales, demonstrated slightly higher self-informant correspondence, with a median r of 0.47.

Although informant reports of both normal and disordered personality have typically shown moderate agreement with self-report, this congruence is far from perfect. Research into the discrepancies between self and informant has frequently focused on imperfect social perception by observers, and attempted to find moderators of agreement inherent to individuals, such as assumed similarity with oneself (Kenny, 1994; Ready, Clark, Watson, & Westerhouse, 2000), or as a function of the relationship, such as level of acquaintanceship (Funder et al., 1995; Funder & Colvin, 1988).

For the most part, discrepancy between self and peers has been treated simply as unwanted error to be minimized, rather than examining the content of disagreement. McCrae, Stone, Fagan, and Costa (1998) did examine the reasons for discrepancies between self-report and spouse ratings, by asking the participants why the disagreement occurred. The researchers found that the most common reason for disagreement was the idiosyncratic interpretation of words or items, which has long been known to introduce error when comparing ratings across individuals (McCrae et al., 1998; Bartoshuk, 2002). McCrae et al. also found that a substantial number of discrepancies arose as a result of the spouse's unawareness of covert feelings and attitudes, and other "enduring features of the judge's understanding and rating of the target" (p. 303).

Such findings suggest that some of the self-other discrepancy reflects a systematic difference between the perception by oneself and by observers. After all, a person has the ability to view his or her own behavior and motivations across different times and settings, making him or her privy to far more information than any observer could ever obtain (Mount et al., 1994). The flip side of this is the "Fish-and-Water Effect" (Kolar et al., 1996). That is, persons become so used to their own behavior that they tune it out, while observers are more aware of these behavioral patterns and therefore make more accurate personality judgments.

To better understand the systematic differences which arise between self and other perception of personality, the current study examines the content of discrepant items between self and peers in a large non-clinical sample. In other words, if peers reliably describe an individual in a consistent way, but the individual does not endorse these traits, how *does* he describe himself? Similarly, if an individual self-reports certain traits, but peers do not agree, what are the peers perceiving? By investigating this incongruence, rather than dismissing it as measurement error, we may gain a better understanding of both normal and disordered personality processes.

2. Methods

2.1. Participants

Participants ($N = 2013$, 1250 male, 763 female) were military recruits at Lackland Air Force Base. The participants in our sample were enlisted personnel, who would eventually receive assignments as military police, mechanics, computer technicians, or other supportive roles. Their mean age was 20 years ($SD = 5$), mean IQ was 104, and 99% were high school graduates. Sixty-four percent described themselves as white, 17% as black, 3% as Asian, 3% as biracial, 1% as Native American, and 11% as another racial group.

Air Force recruits undergo mandatory psychological screenings before beginning basic training, in order to screen out those with Axis I psychopathology. These screenings, however, were not designed to detect or screen out those with Axis II PDs. We administered the Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl, Blum, & Zimmerman, 1997) to a subsample of the participants ($N = 432$). Of those who were interviewed, 19% qualified for at least one definite PD, based on criteria from the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV; American Psychiatric Association, 1994). Another 11% were one criterion short of a diagnosis. On this basis, we estimate the prevalence of PDs in the total sample as 9.4%, roughly equivalent to that found in other non-clinical samples (e.g., Lenzenweger, Loranger, Korfine, & Neff, 1997; Weissman, 1993).

The participants were members of 50 “flights,” groups of 35–52 recruits who had gone through six weeks of basic training together. Fourteen of these flights were single-sex male flights, four were single-sex female flights, and 32 were mixed-gender flights. Recruits in a given flight spend almost all of their time together and get to know each other quite well. Each of the 2013 participants acted as both a nominator and a potential nominee (“target”) in the peer nomination process.

2.2. Materials

Each participant was administered a computerized battery of measures, including the Schedule for Nonadaptive and Adaptive Personality (SNAP), and a peer-report version of the DSM-IV PD criteria.

The SNAP (Clark, 1993) is a factor-analytically derived self-report questionnaire composed of 375 true/false items. Designed as an instrument to investigate dimensional aspects of PD, it includes six validity scales, 15 trait/temperament scales, and ten scales corresponding to the DSM-

III-R PDs. The SNAP diagnostic scales have demonstrated high internal consistency (median values range from 0.72 to 0.83) and validity (Clark, 1993). For this study, only the 10 PD scales were utilized. Endorsing an item in the scored direction added one point to the appropriate scale of the number of items endorsed. In the current study, categorical diagnostic determinations were not made.

The peer-report section of the battery consisted of 106 items, 81 of which were lay translations of the 10 DSM-IV PD criteria.¹ These PD items were constructed by translating the DSM-IV criterion sets for PDs into lay language; resulting items were then reviewed and revised by expert consultants, including a member of the DSM-IV PDs Workgroup.² In a previous study (Thomas, Turkheimer, & Oltmanns, 2003), factor analysis of these peer-report items demonstrated high correspondence (congruence coefficients ranged from 0.87 to 0.97) with factor patterns of self-report models of PDs. Twenty-three filler items were also included in the peer-report measures, based on additional, mostly positive, characteristics, such as “trustworthy and reliable.” Two items, “Please select the people who are your close friends” and “Please select the people who you don’t know at all” were included to investigate how well-acquainted participants were with the individuals they nominated.

Peer-report items were presented to participants in a quasi-random order. For each item, the participant was asked to nominate one or more members of the flight who best exhibit the characteristic in question. For each nomination, the participant assigned the nominee a score (1, 2, or 3), indicating that the nominee “sometimes,” “often,” or “always” displays the characteristic.

The SNAP diagnostic scale scores consisted of the total number of items in the scale endorsed in the correct direction. The peer-report diagnostic scales were calculated by adding up the total number of nominations received for the items in each scale, weighted by the score assigned by each nominator (i.e., 1, 2, or 3), and then dividing by the total number of possible nominators in the flight. The data for each recruit, therefore, consisted of two sets of PD scales: 10 scales of PDs based on self-report, and 10 based on peer nominations.

2.3. *Statistical analysis*

In order to minimize capitalization on chance in our analysis of self- and peer-reported traits, the full sample ($N = 2013$) was randomly divided into two smaller sets: an experimental sample ($N = 1375$) and a cross-validation sample ($N = 638$). Supplemental items were determined using the experimental sample, and then verified using the cross-validation sample.

¹ Although there are 79 total DSM-IV PD criteria, in our process of translating them into lay language, we split two criteria into separate items for the sake of clarity. The first criterion we split was the schizotypal criterion “Inappropriate or constricted affect,” which we rewrote as the items “Shows emotional responses that seem strange or ‘out of sync’” and “Is cold; doesn’t show any feelings.” The second criterion split was the narcissistic item “Is often envious of others or believes that others are envious of him or her,” which we rewrote as “Is jealous of other people” and “Thinks other people are jealous of him or her.” Further, it should be noted that schizoid and schizotypal PDs share two nearly identical DSM criteria. Rather than include the same questions twice, the items “Is cold; doesn’t show any feelings” and “Has no close friends (other than family members)” were included on both the schizoid and schizotypal scales.

² We are grateful to Irving Gottesman and Roger Blashfield for their invaluable assistance with the lay translations of DSM-IV items.

The first step was designed to determine the self-report items which best predict each peer-reported PD scale, over and above the corresponding self-report scale. Using the experimental sample, regressions were performed predicting peer diagnostic scales from the corresponding SNAP scales. For instance, the self-report paranoid scale was used to predict the peer-report paranoid scale. The residuals from this regression represent the unexplained variance once the corresponding self-report scale has been partialled out. These residuals were then correlated, using semipartial correlations, with the entire set of SNAP items, to find the SNAP items which explained the greatest amount of variance in the peer-report scale, over and above the corresponding SNAP scale. The 10 SNAP items which correlated most highly with these residuals were designated the supplemental items for the appropriate diagnostic scale.

The same process was repeated, predicting SNAP scales from the peer scales, then correlating all peer-report items with the residuals of this regression. This likewise established the peer items which best predicted each self-report scale, over and above the corresponding peer scale. There were, therefore, two new sets of items developed: 10 self-report supplemental items, and 10 peer-report supplemental items, each corresponding to a particular type of PD.

This method of determining supplemental items capitalizes on chance. To decrease the effects of chance, further analyses were performed using the cross-validation sample. A unit-weighted supplemental item scale was created by summing scores on the 10 items. In the cross-validation sample, regressions were again performed predicting peer diagnostic scales from the corresponding SNAP PD scale. Then a second set of regressions were performed, predicting the peer scale using both the SNAP PD scale and the corresponding SNAP supplemental items scale as determined in the experimental sample, to investigate how much variance the supplemental items explained, over and above the standard SNAP scales. This same method was used to investigate the peer supplemental items, predicting SNAP scale scores within the cross-validation sample using the corresponding peer PD scale, first alone, and then with the addition of the appropriate supplemental items scale.

3. Results

The overall correspondence between self- and peer-report in our study were somewhat smaller than that demonstrated in previous studies. In large part, this is because our analyses are based on a cross-validation sample, which greatly reduces the effects of chance. The use of cross-validation is a stringent methodology that will probably result in smaller concordance between self and peer. The lower self–peer correspondence in the present study does not appear to be due to unreliable reporting or poorly acquainted peers. Reliability of peer nominations was quite high, with mean intraclass correlations across flights ranging from 0.88 to 0.97.

3.1. Regression results

The regressions predicting either self-reported or peer-reported PDs using only the corresponding diagnosis of the other scale explained little of the variance in the cross-validation sample. As summarized in Table 1, the correlation between the same scales assessed by self- and peer-report was low, ranging from 0.11 to 0.22.

Table 1

Correlation between self-report and peer-report with and without self-report supplemental items in the cross-validation sample ($N = 688$)

Diagnostic scale	Correlation	Including self-report supplemental items	Content of self-report supplemental items
Paranoid	0.11	0.26**	Angry, hostile
Schizoid	0.14	0.21**	Hostile, mistrustful of others
Schizotypal	0.18	0.22**	Odd, different from others
Antisocial	0.17	0.30**	Hostile, angry, extraverted
Borderline	0.14	0.15	Hostile, pessimistic, self-harm
Histrionic	0.22	0.31**	Outgoing, gregarious, hostile
Narcissistic	0.15	0.31**	Outgoing, gregarious, energetic
Avoidant	0.12	0.17*	Pessimistic, poor self-image
Dependent	0.16	0.19*	Poor self-image, pessimistic
Obsess–Comp	0.14	0.27**	Gregarious, hard-working

* $p < 0.01$.

** $p < 0.001$.

Predicting peer-reported PDs using both the corresponding SNAP diagnostic scale and the appropriate supplemental items increased predictive ability somewhat, with r ranging from 0.15 to 0.31 in the cross-validation sample. These results are summarized in Table 1.

Similarly, using both the peer diagnostic scale and the appropriate supplemental items to predict SNAP PDs increased the predictive ability, with r ranging from 0.14 to 0.37 in the cross-validation sample, as seen in Table 2. Adding the peer-report supplemental items increased predictive ability the most for Cluster B disorders (improvement in r ranged from 0.05 to 0.13) and the least for Cluster C PDs (improvement ranged from 0.00 to 0.07).

The individual self-report supplemental items, defined as the 10 highest correlating SNAP items with the peer measure residuals for each diagnostic category, demonstrated semipartial

Table 2

Correlation between self-report and peer-report with and without peer-report supplemental items in the cross-validation sample ($N = 688$)

Diagnostic scale	Correlation	Including peer supplemental items	Content of peer supplemental items
Paranoid	0.11	0.13	Cold, unfeeling
Schizoid	0.14	0.20**	Aloof, loner
Schizotypal	0.18	0.18	Loner, poor self-image
Antisocial	0.17	0.29**	Reckless, troublemaker, likeable
Borderline	0.14	0.17*	Uncaring, unfriendly, unreliable
Histrionic	0.22	0.35**	Likeable, outgoing, many friends
Narcissistic	0.15	0.25**	Reckless, insensitive, likeable
Avoidant	0.12	0.23**	Meek, unlikable, few friends
Dependent	0.16	0.20**	Meek, needy
Obsess–Comp	0.14	0.15	Level-headed, unexciting

* $p < 0.01$.

** $p < 0.001$.

correlations (r_p) ranging from 0.10 to 0.22 in the experimental sample. Several supplemental items were negatively correlated with the corresponding scale. Negative semipartial correlations for supplemental self-report items indicate that the participant tended not to endorse these traits. The highest single correlating supplemental item was with Antisocial PD (Item #21, “I often quarrel with others,” $r_p = 0.22$). In other words, individuals whose peers saw them as antisocial were likely to endorse this item. Other highly correlated items included paranoid PD (Item #70, “Some people consider me hostile,” $r_p = 0.20$), and avoidant PD (Item #237, “My future looks very bright to me,” $r_p = -0.19$). The 10 items comprising each self-report supplemental item scale are described qualitatively in Table 1, and listed in Table 3.

The peer supplemental items, defined as the 10 peer-report items which correlated most highly with the SNAP residuals for each diagnostic category, had semipartial correlations ranging from 0.07 to 0.19 in the experimental sample. For supplemental peer items, negative semipartial correlations indicate that the individual tended not to be nominated by his or her peers for this item. The highest single correlating supplemental peer item was a negative correlation with histrionic PD (Item #15, “Is reserved or shy when meeting new people because he/she feels inadequate (not as good as other people),” $r_p = -0.19$). In other words, people who described themselves (on the SNAP) as being histrionic were unlikely to be nominated for this item by their peers. Other highly correlated items included antisocial PD (Item #68, “Repeatedly gets in trouble with the police,” $r_p = 0.18$) and paranoid PD (Item #5, “Is trustworthy and reliable,” $r_p = -0.17$). The 10 highest correlated peer items for each diagnostic category are summarized in Table 2, and listed in Table 4.

4. Discussion

The correspondence between self-report and peer nomination of PDs in our study was fairly low. Knowing a participant’s score on a SNAP PD scale imparts limited predictive ability about his or her peer-reported score, and vice-versa. In other words, there is a rift between how participants describe themselves and how others describe them. By examining the self-report supplemental items, however, we gain additional information regarding how individuals view themselves. And, conversely, by investigating the peer-report supplemental items, we can better understand how their peers view them.

The supplemental items are items in one domain (self or peer) that correlate with the scale for a given disorder in the other domain, over and above the corresponding scale. In examining the supplemental items, we discovered several different ways of describing their content. Some are items which were included in the original corresponding scale, but which are such good predictors that they remain in the model after the total score on the original scale has been partialled out. An example of this is the antisocial PD (ASP) peer item “Repeatedly gets in trouble with the police.” Although it is included in the ASP scale, this item is such a strong predictor of self-reported ASP that it correlates over and above the original peer scale.

A second variation of supplemental items includes items that are similar in content to the corresponding PD scale, but which were not included in the original scale. The peer supplemental items for schizoid PD consist of negatively correlated histrionic and narcissism items. For example, a negative correlation with the item “Is unhappy when he/she is not the center of

Table 3

Summary of the 10 most highly correlated supplemental self-report items for each peer-report diagnostic category ($N = 1375$)

Peer	Self-report (SNAP) items									
Paranoid	Some consider me hostile	Often quarrel with others	Temper gets me in trouble	Voice isn't as expressive as others'	Was often sent to the principal for fighting	<i>Trouble throwing things away</i>	<i>Very hard to make me angry</i>	Become angry more easily than most	<i>No trouble controlling my anger</i>	Anger frequently gets the better of me
<i>R</i>	0.20	0.19	0.16	0.14	0.13	-0.13	-0.13	0.13	-0.12	0.12
Schizoid	People think I'm odd	Often quarrel with others	I have left my body	Been told my behavior is strange	Some consider me hostile	Others say they can't understand me	<i>My future looks bright</i>	People try to make me look foolish	I never say anything bad about others	<i>No trouble controlling my anger</i>
<i>R</i>	0.16	0.15	0.13	0.13	0.12	0.12	-0.12	0.11	0.11	-0.11
Schizotypal	<i>My future looks bright</i>	People think I'm odd	I never say anything bad about others	<i>Trouble throwing things away</i>	<i>People don't care enough about manners</i>	Get so upset I feel like hurting myself	<i>Sometimes drive above speed limit</i>	<i>I'm self-reliant</i>	Often quarrel with others	Can't stand being alone
<i>R</i>	-0.14	0.14	0.12	-0.12	-0.11	0.11	-0.11	-0.10	0.10	0.10
Antisocial	Often quarrel with others	Some consider me hostile	Temper gets me in trouble	Really enjoy speaking in public	<i>Trouble throwing things away</i>	I anger quickly if criticized even a little	<i>Very hard to make me angry</i>	<i>Fear of criticism holds me back</i>	Others can't keep up with my pace	Really enjoy performing in public
<i>R</i>	0.22	0.15	0.13	0.12	-0.12	0.12	-0.11	-0.10	0.10	0.11
Borderline	<i>My future looks bright</i>	Some consider me hostile	I have made a mess of my life	<i>Trouble throwing things away</i>	Never cheated on a test	Often quarrel with others	When I go out, like to know what I'm doing	Try suicide is only way to be taken seriously	Hurt myself on purpose frequently	<i>I'm self-reliant</i>
<i>R</i>	-0.15	0.12	0.11	-0.11	0.11	0.10	0.10	0.10	0.10	-0.10
Histrionic	Really enjoy public speaking	Often quarrel with others	Some consider me hostile	If relationship ends, new one right away	<i>Voice isn't as expressive as others'</i>	Those I care about have been unfaithful	<i>I'm reserved around others</i>	Temper gets me in trouble	My best friends are now worst enemies	I get others to do my work
<i>R</i>	0.15	0.14	0.13	0.13	-0.12	0.12	-0.12	0.11	0.11	0.11
Narcissistic	<i>Voice isn't as expressive as others'</i>	<i>Fear of criticism holds me back</i>	Really enjoy public speaking	<i>Around others I keep to myself</i>	My pace is quick and lively	<i>I'm no good at flirting</i>	Often quarrel with others	I'm an energetic person	Others can't keep up with my pace	Some consider me hostile
<i>R</i>	-0.18	-0.17	0.16	-0.15	0.14	-0.14	0.14	0.13	0.13	0.12

Table 3 (continued)

Peer	Self-report (SNAP) items									
Avoidant	<i>My future looks bright</i>	Let others do things so done right	People think I'm odd	<i>I'm self reliant</i>	<i>I handle my life pretty well</i>	I have made a mess of my life	Only do enough just to get by	<i>I like making my own decisions</i>	Hate being alone for any length of time	Others tell me they can't understand me
<i>R</i>	–0.19	0.15	0.15	–0.14	–0.13	0.12	0.12	–0.12	0.12	0.11
Dependent	Let others do things so done right	<i>My future looks bright</i>	I have made a mess of my life	<i>People don't care enough about manners</i>	People try to make me look foolish	Others tell me they can't understand me	People think I'm odd	<i>Sometimes drive above speed limit</i>	Often quarrel with others	Only do enough just to get by
<i>R</i>	0.16	–0.15	0.12	–0.12	0.12	0.11	0.11	–0.11	0.10	0.10
Obsessive–compulsive	<i>Voice isn't as expressive as others'</i>	<i>Trouble throwing things away</i>	Really enjoy public speaking	Others can't keep up with my pace	<i>Fear of criticism holds me back</i>	<i>Never throw out anything I might need</i>	Others say I drive myself hard	<i>Uncomfortable unless I know people well</i>	Enjoy working hard	My pace is quick and lively
<i>R</i>	–0.15	–0.15	0.14	0.13	–0.11	–0.11	0.10	–0.10	0.10	0.10

Note: Negatively correlated items are written in italics.

Table 4

Summary of the 10 most highly correlated supplemental peer-report items for each self-report diagnostic category ($N = 1375$)

Self	Peer-report items									
Paranoid	<i>Trustworthy and reliable</i>	<i>Cheerful and optimistic</i>	<i>Pursues tasks diligently</i>	<i>Remains calm under stress</i>	<i>Is sincere and genuine</i>	<i>Seems to be under-standing</i>	<i>Strict and rigid morals</i>	<i>Charismatic and leadership</i>	<i>Compassionate and concerned</i>	<i>Acts as a leader</i>
R	-0.17	-0.16	-0.16	-0.14	-0.14	-0.14	-0.13	-0.13	-0.13	-0.12
Schizoid	<i>Needs others to admire him/her</i>	<i>Unhappy when not the center of attention</i>	<i>Good sense of humor</i>	<i>Uses appearance to get attention</i>	<i>Exaggerated emotions</i>	<i>Your close friends</i>	<i>Dominant in relationships</i>	<i>Thinks about power or love</i>	<i>Cheerful and optimistic</i>	<i>Thinks others are jealous</i>
R	-0.15	-0.14	-0.14	-0.14	-0.13	-0.13	-0.13	-0.13	-0.13	-0.13
Schizotypal	<i>Cheerful and optimistic</i>	<i>Trustworthy and reliable</i>	<i>Pursues tasks diligently</i>	<i>Is sincere and genuine</i>	<i>Seems to be under-standing</i>	<i>Your close friends</i>	<i>Charismatic and leadership</i>	<i>Assertive in interactions</i>	<i>Remains calm under stress</i>	<i>Acts as a leader</i>
R	-0.14	-0.13	-0.12	-0.11	-0.11	-0.11	-0.10	-0.10	-0.10	-0.09
Antisocial	<i>Gets in trouble with the police</i>	<i>Reckless lifestyle</i>	<i>Does things b/c others won't do right</i>	<i>Strict and rigid morals</i>	<i>Good sense of humor</i>	<i>Is generally a tender person</i>	<i>Needs such a perfect job that doesn't finish</i>	<i>Is very rigid and stubborn</i>	<i>Too concerned about details and lists</i>	<i>Works so much that never has fun</i>
R	0.18	0.15	-0.15	-0.15	0.14	-0.13	-0.13	-0.12	-0.11	-0.11
Borderline	<i>Is generally a tender person</i>	<i>Strict and rigid morals</i>	<i>Pursues tasks diligently</i>	<i>Trustworthy and reliable</i>	<i>Is sincere and genuine</i>	<i>Seems to be under-standing</i>	<i>Sympathetic & kind to others</i>	<i>Remains calm under stress</i>	<i>Compassionate and concerned</i>	<i>Sensitive to others' needs</i>
R	-0.16	-0.15	-0.15	-0.14	-0.14	-0.14	-0.13	-0.13	-0.13	-0.13
Histrionic	<i>Shy around people b/c feels inadequate</i>	<i>Not interested in close relationships</i>	<i>Good sense of humor</i>	<i>No close friends</i>	<i>Avoids teams, b/c fears rejection</i>	<i>Nervous b/c doesn't trust others</i>	<i>Prefers to do things alone</i>	<i>Is generally a tender person</i>	<i>Talks in a vague way that lacks detail</i>	<i>Inhibited with friends b/c fears ridicule</i>
R	-0.19	-0.19	0.18	-0.18	-0.16	-0.16	-0.16	-0.14	-0.14	-0.14
Narcissistic	<i>Is generally a tender person</i>	<i>Good sense of humor</i>	<i>Sympathetic & kind to others</i>	<i>Reckless lifestyle</i>	<i>Seems to be under-standing</i>	<i>Creativity or artistic talent</i>	<i>Is irresponsible</i>	<i>Gets in trouble with the police</i>	<i>Sensitive to the needs of others</i>	<i>Is sincere and genuine</i>

Table 4 (continued)

Self	Peer-report items									
R Avoidant	-0.14	0.12	-0.12	0.12	-0.11	0.11	0.11	0.11	-0.10	-0.10
	<i>Has a good sense of humor</i>	<i>Cheerful and optimistic</i>	<i>Your close friends</i>	<i>Needs others to admire him</i>	<i>Uses appearance to get attention</i>	<i>Is a talented athlete</i>	<i>Unhappy when not the center of attention</i>	<i>Assertive in inter-actions</i>	<i>Exaggerated emotions</i>	<i>Dominant in relationships</i>
R Dependent	-0.17	-0.16	-0.16	-0.15	-0.15	-0.15	-0.14	-0.14	-0.13	-0.13
	<i>Doesn't care about praise or criticism</i>	<i>Seems to be independent</i>	<i>Does things b/c others won't do right</i>	<i>Is very rigid and stubborn</i>	<i>Doesn't show any feelings</i>	<i>Willing to take a stand for beliefs</i>	<i>Not concerned about others' feelings</i>	<i>Is unmoved after hurting someone</i>	<i>Dominant in relationships</i>	<i>Articulate and persuasive</i>
R Obsessive-compulsive	-0.14	-0.13	-0.12	-0.11	-0.11	-0.11	-0.10	-0.10	-0.09	-0.09
	<i>Reckless lifestyle</i>	<i>After a break up, quickly finds another</i>	<i>Reckless lack of concern for safety</i>	<i>Lies to people, or cons people</i>	<i>Gets in trouble with the police</i>	<i>Inappropriate sexually seductive</i>	<i>Unhappy when not the center of attention</i>	<i>Needs others to admire him/her</i>	<i>Uses appearance to get attention</i>	<i>Pursues tasks diligently</i>
R	-0.11	-0.09	-0.09	-0.09	-0.09	-0.08	-0.08	-0.07	-0.07	0.07

Note: Negatively correlated items are written in italics.

attention” accurately describes the prototypical schizoid PD, but was not a part of the original scale. Similarly, the self-report obsessive–compulsive PD (OCPD) supplemental items include “Other people sometimes have trouble keeping up with the pace I set.” Although this item might describe the work ethic of OCPD, it was actually a part of the SNAP’s positive temperament scale, rather than the original OCPD self-report scale.

Another possibility for supplemental items seems to be unexplained error variance. These items do not show an easily identifiable common theme. An example of this can be seen in the self-report schizotypal supplemental items, which include items ranging from “I never say anything bad about people, even if they’re not around to hear” (from the Rare Virtues scale of the SNAP) to a negative correlation with the item “People today aren’t concerned enough about good manners.” Although this assortment of supplemental items may reflect the idiosyncratic thought processes of people with schizotypal PD, it is difficult to find a common theme to the inclusion of the items.

Yet another type of supplemental item is an example of what test construction literature refers to as a “subtle item.” These are self-reported items which are less obvious, non-pejorative examples of PD traits. For example, participants who were described by peers as narcissistic tended to self-report being extremely outgoing, gregarious, and likeable. In other words, the participants self-reported positive traits, subtly displaying the very narcissism described by peers.

Finally, for some items there appears to be systematic differences between self and other in the perception of a trait. For example, knowing that peers describe a given participant as high on paranoid PD provides only a small estimate of the way that the participant describes him or herself on the SNAP paranoid PD scale ($r = 0.11$). By examining self-report supplemental paranoid items (i.e., the self-report items which correlate with the peer paranoid scale, over and above the self-reported paranoid items), we begin to understand some elements of this discrepancy. We find that the self-report supplemental paranoid items suggest anger and hostility. This may indicate several things. One possibility may be that the individual behaves in a paranoid manner, but is either unaware of this or unwilling to admit to it, and instead describes himself as hostile and angry. Another possibility is that the individual really does feel hostile and angry, but peers, having only overt behavior on which to judge, interpret his hostility as being based on paranoia.

This analysis of supplemental items can also be turned around to investigate the peer descriptions of self-described paranoid participants. Again, knowing that a participant describes himself as high on the paranoid scale of the SNAP provides only a partial indication of the way that his peers describe him ($r = 0.11$). But, by examining the peer supplemental items that predict self-reported paranoia over and above the peer paranoid scale, we gain additional understanding of the way in which individuals who describe themselves as paranoid are perceived by their peers. The most highly correlated peer supplemental items include “Is trustworthy and reliable” ($r_p = -0.17$), “Is cheerful and optimistic” ($r_p = -0.16$), “Is sincere and genuine” ($r_p = -0.14$), and other items indicating that the individual is perceived as cold and detached. Again, this may suggest that the individual is cold and detached (as peers describe him), but that the individual instead endorses items indicative of paranoia. Alternatively, it may be that the individual does, indeed, feel paranoid as he describes, and that this paranoia leads him to keep others at a distance. Without access to the individual’s internal motivation, his or her peers may ascribe it to the individual being cold or aloof, rather than paranoid.

A possible limitation of the current study should be noted. The sample of participants in this study, Air Force recruits undergoing basic training, is not typical of most psychological literature. Further, their relatively young age ($M = 20$ years) may limit their psychological insight, which could contribute to decreased self–peer agreement. Although we would suggest that such a sample is no less representative of the general population than the college undergraduates who are the focus of much research, the generalizability of our results to other populations remains to be seen.

Our study seeks to understand the relatively large discrepancy between self-report and peer-report of PDs. One possibility is that self-reports may be fraught with defensiveness and attempts, both conscious and unconscious, to respond in a socially desirable manner. While this is certainly a possibility, it seems unlikely that it accounts for the entire discrepancy. Past studies have shown that, while self- and peer-reports do tend to differ qualitatively, the amount of pathology elicited by self-report tends to be equal to or greater than that from peers (Klonsky et al., 2002; Riso, Klein, Anderson, Ouimette, & Lizardi, 1994). Further, the participants in our study who were nominated by peers as personality disordered did not, for the most part, deny negative traits, but rather endorsed other types of pathology. On the other hand, those who were peer nominated as having narcissistic, histrionic, or antisocial traits did tend to endorse positive traits, and described themselves as outgoing and flirtatious, as seen in Table 3. Those peer-nominated as narcissistic, especially, gave ringing endorsements of themselves, describing themselves as hard-working, outgoing, and personable. In each of these cases, however, the discrepancy does not seem to be due to defensiveness or desire to appear socially desirable, but rather to a genuine belief in one's own superiority. In other words, what may at first appear to be defensiveness is instead simply another expression of the PD traits.

Another explanation for the self–peer discrepancy is that participants are accurately reporting their personality, which they base on both overt behavior and internal states. Peers, on the other hand, can make judgments based only on overt behavior, which leads to misinterpretation of the behavior, albeit a consistent one across peers. Numerous studies have shown that, in the domain of normal personality, there is greater correspondence between peer- and self-ratings for traits which are more easily observable or public (Lanning & Gough, 1991; Funder, 1980; Funder & Dobroth, 1987; Kenrick & Stringfield, 1980). For traits that are less overtly perceptible, peers may make judgments based on their understanding of their own personalities, resulting in decreased self–peer congruence (Ready et al., 2000). This explanation is consistent with the higher correlations for disorders with more overt behavioral traits, such as antisocial and histrionic, than for more internalized disorders, like obsessive–compulsive and avoidant. It is simply easier for peers to accurately identify more overt traits, such as the dramatic tendencies of histrionic PD, while they might have more difficulty in recognizing less dramatic traits, such as social anxiety or perfectionism. In addition, even when peers accurately identify behaviors, they may have difficulty in ascribing the underlying reason for the behavior.

Neither self-report nor peer nomination is a perfect measure of the individual's personality (Clark, Livesley, & Morey, 1997). Each filters reality through a different lens. Peers have access only to overt behaviors, not the internal state of the target. Self-report, although able to incorporate covert behaviors and greater amounts of information, is hampered by a lack of objectivity, an especially large problem in people with PDs. The lack of convergence between self- and peer-report in personality research might be better understood as a byproduct of personality itself, rather than as a limitation of the tools of measurement. Rather than focusing on the low cor-

respondence between self and other, the use of supplemental items may answer a more fundamental question. If peers see us differently than we see ourselves, how *do* they see us? The results of this study suggest that the use of supplemental items may provide a qualitative and systematic answer to this question.

Acknowledgements

The authors would like to express their great appreciation to Edna Fiedler for her many contributions to this study. Thanks also go to Pat Kyllonen, Janice Hereford, and Sgt. Randy John at Lackland Air Force Base, for their aid in testing participants and analyzing data. We would also like to thank Lee Anna Clark for her suggestions and comments on an earlier draft of this paper. This study was supported by a grant from the National Institute of Mental Health (MH51187).

References

- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- Asher, S. R., & Hymel, S. (1981). Children's social competence in peer relations: sociometric and behavioral assessment. In J. D. Wine & M. D. Syne (Eds.), *Social competence* (pp. 125–157). New Mexico: Guilford.
- Bartoshuk, L. (2002). Self reports and across-group comparisons: a way out of the box. *APS Observer*, 15(7), 26–28.
- Clark, L. A. (1993). *Schedule for nonadaptive and adaptive personality (SNAP)*. Minneapolis, MN: University of Minnesota Press.
- Clark, L. A., Livesley, W. J., & Morey, L. (1997). Personality disorder assessment: the challenge of construct validity. *Journal of Personality Disorders*, 11, 205–231.
- Fiedler, E. R., Oltmanns, T. F., & Turkheimer, E. (in press). Personality problems and adjustment to military life: predictive validity of self and peer reports. *Military Medicine*.
- Fuhrman, R. W., & Funder, D. C. (1995). Convergence between self and peer in the response-time processing of trait-relevant information. *Journal of Personality and Social Psychology*, 69, 961–974.
- Funder, D. C. (1980). On seeing ourselves as others see us: self-other agreement and discrepancy in personality ratings. *Journal of Personality*, 48, 473–493.
- Funder, D. C. (1991). Judgments of personality: a new approach to the accuracy issue. In D. J. Ozer, & J. M. Healy (Eds.), *Perspectives in personality* (vol. 3, Part B, pp. 107–132). London, England: Jessica Kingsley Publishers, Ltd.
- Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology*, 55, 149–158.
- Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: properties associated with interjudge agreement. *Journal of Personality and Social Psychology*, 52, 409–418.
- Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: interpersonal relations, similarity, and acquaintanceship. *Journal of Personality and Social Psychology*, 69, 656–672.
- Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43–62.
- Hurley, J. R. (1988). Interpersonal correlates of individuals' discrepant ratings by self and peers. *The Journal of Social Psychology*, 128, 653–665.
- Kane, J. S., & Lawler, E. E., III (1978). Methods of peer assessment. *Psychological Bulletin*, 85, 555–586.
- Kenny, D. A. (1994). *Interpersonal perception: a social relations analysis*. New York: Guilford Press.
- Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: lessons from the person-situation debate. *American Psychologist*, 43, 23–34.

- Kenrick, D. T., & Stringfield, D. O. (1980). Personality traits and the eye of the beholder: crossing some traditional philosophical boundaries in the search for consistency in all of the people. *Psychological Review*, *87*, 88–104.
- Klonsky, E. D., Oltmanns, T. F., & Turkheimer, E. (2002). Informant reports of personality disorder: relation to self-report, and future research directions. *Clinical Psychology: Science and Practice*, *9*, 300–311.
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality*, *64*, 311–337.
- Lanning, K., & Gough, H. G. (1991). Shared variance in the California Psychological Inventory and the California Q-Set. *Journal of Personality and Social Psychology*, *60*, 596–606.
- Lenzenweger, M. F., Loranger, A., Korfine, L., & Neff, C. (1997). Detecting personality disorders in a nonclinical population: application of a 2-stage for case identification. *Archives of General Psychiatry*, *54*, 345–351.
- Levesque, M. J., & Kenny, D. A. (1993). Accuracy of behavioral predictions at zero acquaintance: a social relations analysis. *Journal of Personality and Social Psychology*, *65*, 1178–1187.
- Maassen, G. H., Goossens, F. A., & Bokhorst, J. (1998). Ratings as validation of sociometric status determined by nominations in longitudinal research. *Social Behavior and Personality*, *26*, 259–274.
- McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81–90.
- McCrae, R. R., Stone, S. V., Fagan, P. J., & Costa, P. T., Jr. (1998). Identifying causes of disagreement between self-reports and spouse ratings of personality. *Journal of Personality*, *66*, 285–313.
- Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology*, *79*, 272–280.
- Oltmanns, T. F., Turkheimer, E., & Strauss, M. E. (1998). Peer assessment of personality traits and pathology in female college students. *Assessment*, *5*, 53–65.
- Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured interview for DSM-IV personality (SIDP-IV)*. Washington, DC: American Psychiatric Press.
- Putallaz, M., & Gottman, J. M. (1981). Social skills and group acceptance. In S. R. Asher & J. M. Gottman (Eds.), *The development of children's friendships* (pp. 116–149). Cambridge: Cambridge University Press.
- Ready, R. E., Clark, L. A., Watson, D., & Westerhouse, K. (2000). Self- and peer-reported personality: agreement, trait ratability, and the “self-based heuristic.” *Journal of Research in Personality*, *34*, 208–224.
- Riso, L. P., Klein, D. N., Anderson, R. L., Ouimette, P. C., & Lizardi, H. (1994). Concordance between patients and informants on the personality disorder examination. *American Journal of Psychiatry*, *151*, 568–573.
- Schwarzwald, J., Koslowsky, M., & Mager-Bibi, T. (1999). Peer ratings versus peer nominations during training as predictors of actual performance criteria. *Journal of Applied Behavioral Science*, *35*, 360–372.
- Thomas, R. C., Turkheimer, E., & Oltmanns, T. F. (2003). Factorial structure of pathological personality as evaluated by peers. *Journal of Abnormal Psychology*, *112*, 81–91.
- Weissman, M. (1993). The epidemiology of personality disorders: a 1990 update. *Journal of Personality Disorders* (Suppl. 1), 44–62.
- Westen, D. (1997). Divergences between clinical and research methods for assessing personality disorders; implications for research and the evolution of Axis II. *American Journal of Psychiatry*, *154*, 895–903.
- Zimmerman, M., Pfohl, B., Stangl, D., & Corenthal, C. (1986). Assessment of DSM-III personality disorders: the importance of interviewing an informant. *Journal of Clinical Psychiatry*, *47*, 261–263.