MCMI-II predictors of judgmental tendencies as measured by the first impression interaction procedure

Alan R. King*, Allison N. Pate

Psychology Department, University of North Dakota, PO Box 8380, Grand Forks, ND 58202-8380, USA

Received 19 November 2001; received in revised form 23 April 2002; accepted 14 June 2002

Abstract

The present study investigates the validity of the MCMI-II personality disorder scales to predict judgmental tendencies as measured by (King & Pate, 2002). A total of 137 women and 48 men completed two separate 25-min uncontrolled interaction sessions as specified by the FIIP. Scores on the 14-item FIIP rating form were then combined to quantify first impression central tendency (Judgment Index), variability (Judgment Variability Index), and the degree to which participants judged others more harshly than self (Judgment Ratio). Judgmental tendencies as measured by the FIIP were associated with DSM-IV cluster B (“dramatic, emotional, or erratic”) features among women and cluster A (“odd or eccentric”) qualities among men. These gender by personality interactions were particularly evident for the histrionic and narcissistic scales. Personality features such as social alienation and insecurity provided the best predictors of judgmental tendencies among men, while harsh FIIP judgment ratings were more common among self-confident women who sought attention and affirmation from others. It was speculated that initial social appraisals might be differentially coarsened by personality traits of insecurity in men and grandiosity in women.

© 2002 Elsevier Science Ltd. All rights reserved.

The potential maladaptive consequences of extreme attitudinal and behavioral response tendencies have been discussed in the clinical psychology literature (Millon & Davis, 1996, 1999). Judgmental attitudes and behavior are distinguished by their negative valence and statistical deviance from average evaluation standards. Judgmental tendencies may be often be observed to occur in the absence of situational or environmental provocation. In fact, Cronbach’s (1955) seminal work in person perception culminated with his historic conclusion that the judge was often more important than the interaction in interpersonal impression in formation. The cen-
trality of the judgmental trait in personality pathology is reflected in contemporary DSM-IV (American Psychiatric Association, 1994) conceptualizations of antisocial (“irritability and aggressiveness”), passive-aggressive (“sullen and argumentative”; “unreasonably criticizes and scorns authority”), paranoid (“counterattacks to perceived challenges to character not apparent to others”; “suspects to be exploited, harmed or deceived by others”; “persistently bears grudges”), narcissistic (“often envious of others or believes that others are envious of him or her”), borderline (“inappropriate, intense anger”, “relationships alternate between extremes of idealization and devaluation”), and even obsessive-compulsive (“inflexible about matters of morality, ethics, or values”) disorders. Some empirical evidence has been found for judgmental tendencies among individuals manifesting personality (Matano & Locke, 1995; Sim & Romney, 1990; Thomas-Peter, 1993) and other forms of psychological (Arthur, 1966; Parsonson, 1969) disturbance, but relationships between personality disorder dimensions and judgmental behavior have been generally ignored.

Judgmental behavior has been difficult to investigate systematically since a conventional method for its measurement has not previously been established. One paradigm that has held promise is the measurement of first impressions since these judgments and associated behaviors are largely insulated from the effects of prior exposure and experiences with the target. Most first impression studies (over 150 now available) have examined participant reactions to a small number of unique, precisely-defined targets that are presented indirectly in the form of written vignettes, photographs, or videotaped segments. The direction (positive versus negative) and intensity of first impression ratings have been examined, but qualitative analyses of the adjectives used to describe targets have also been popular. Ratings scales in these studies have tended to be relatively simple and idiosyncratic to purposes of the respective investigation (Bryan, Coleman, Ganong, & Bryan, 1986; Walther, 1993).

King and Pate (2002) recently developed the First Impression Interaction Procedure (FIIP) which is distinguished by its reliance on diverse targets encountered in a more natural social setting as a strategy for generating ecologically valid measures of first impression central tendency, variability, and proclivity to judge others more or less harshly than self. Judgmental tendencies as quantified from the FIIP have the benefit of being derived from samples of actual rating behavior exhibited after diverse target encounters.

One popular method to assess personality disorder dimensions is provided by the Millon Clinical Multiaxial Inventory (MCMI-II). The MCMI-II provides theoretically derived measures which predict diagnostic probabilities and a wide range of cognitive, affective, perceptual, interpersonal, and behavioral clinical attributes (Millon, 1987). MCMI-II interpretive guidelines appear generally consistent with the DSM-III-R personality disorder formulations. The sadistic, self-defeating, and passive-aggressive categories have been subsequently eliminated from the DSM-IV, but the MCMI-II dimensions otherwise measure, to considerable extent, the attributes associated with their respective contemporary DSM-IV personality disorder features. MCMI-II personality dimensions are quantified in base rate (BR) scores which are designed to assure that the prevalence of scale elevations (BR > 75) corresponds closely with the prevalence of the disorder found among psychiatric patients. This scaling strategy is useful in providing information about the probability of particular Axis II diagnoses among individuals presenting for treatment, but the relatively low prevalence of extreme personality features found among members of the general public will result in substantial false positive classification errors if recommended base
rate interpretive guidelines are applied to profiles where the fundamental assumption of extant symptomatology is violated. Millon has responded to these interpretive risks by discouraging the use of the MCMI-II in non-clinical populations, but others (Choca, Shanley, & Van Denburg, 1992) have instead interpreted MCMI-II base rate scores as measures of stylistic personality differences. This perspective assumes that base rate scores provide information about personality attributes distributed throughout the respective dimensions for both adjusted and clinical populations, while supporting Millon’s reservations about the inappropriateness of recommended MCMI-II cutoffs for assessing diagnostic probabilities among adjusted individuals.

The relative merits of theoretically driven versus factor-derived personality scale construction and application remains subject to debate. Personality inventory analysis has traditionally focused on unitary subscale scores and the resulting profile categorization. Others have noted that the inherent covariance in inventory subscales can lead to misinterpretation of personality protocols. Five factor theory (Digman, 1990) has emerged to support the development of inventories such as the NEO-PI (Costa & McCrae, 1992) which derive a smaller subset of relatively independent personality dimensions. Indeed, eight MCMI-II factors were identified during scale construction, but these scores bear their own interpretive challenges and are not easily linked to meaningful criterion variables, such as DSM-IV diagnostic probabilities. Similar concerns have been raised about the usefulness of the NEO-PI in assessing personality disorders (Coolidge, Becker, DiRito, & Durham, 1994). The DSM-IV has noted that personality disorders can often, however, be clustered together into groupings based upon shared features. For example, it is observed (pp. 629–630) that paranoid, schizoid, and schizotypal personalities share tendencies toward “odd or eccentric” behavior (Cluster A). Antisocial, borderline, histrionic, and narcissistic personalities were distinguished by their penchant for “dramatic, emotional, or erratic” actions (Cluster B). Avoidant, dependent, and obsessive-compulsive personalities appear to share anxiety and fearfulness (Cluster C).

This study investigates the validity of the MCMI-II personality disorder scales to predict judgmental tendencies as measured by the FIIP. An attempt is also made to identify personality factors which might predispose judgmental actions as they might occur in everyday life.

1. Method

1.1. Participants

Undergraduate students enrolled in selected courses of personality theories, introductory, abnormal, and clinical psychology at a major Midwestern state university were invited to participate in this study of “interpersonal perceptiveness” for extra credit (about 4% of final grade). Data from 226 (out of 479) participants met the FIIP eligibility requirements for inclusion in the analysis (e.g. completion of two interaction sessions with six unfamiliar partners). This initial attrition (53% of the initial sample) occurred randomly as a function of the absenteeism of, or unexpected familiarity with, one or more of each participant’s assigned interaction partners. A decision was also made to exclude a small number of minority students (n = 20) to control for possible ethnicity effects. An additional 41 participants were excluded due to uninterpretable MCMI-II profiles (Validity > 0; Desirability or Debasement Base Rates > 85; raw Dis-
closeness > 590 or < 145) as determined by guidelines recommended in other sources (Choca et al., 1992; Millon, 1987). The final sample was comprised of 137 women and 48 men between the ages of 18 and 54 (\(M = 24, SD = 7\)) who were generally classified at a sophomore or junior academic level in terms of cumulative college credits (\(M = 69, SD = 30\)).

1.2. Materials

The FIIP assesses participant initial reactions to six unfamiliar partners encountered during two separate 25-min uncontrolled interaction sessions. Scores on the 14-item FIIP rating form are combined to quantify first impression central tendency (Judgment Index), variability (Judgment Variability Index), and the degree to which participants judged others more harshly than self (Judgment Ratio). Items on the FIIP rating scale have clearly distinguished poles in terms of social desirability, and participants are provided an opportunity to rate their interaction partners while also providing self-evaluations and predictions about how they are likely to be rated by the others using the same scale. JI, JVI, and JR scores range from 1 to 7, 0 to 3.0, and 0.14 to 7.0, respectively. A certain level of session to session variability in the FIIP judgment indices is inevitable given uncontrolled partner and interaction situational effects, but King and Pate (2002) found that internal consistency and temporal stability coefficients for combined data from two sessions and six partners ranged from 0.85 to 0.94. Correlations between the FIIP judgment indices were found to range from 0.41 to 0.46. The FIIP provides a method of identifying individuals distinguished by their tendency to form and express harsh first impressions of others (\(\uparrow\)JI, \(\uparrow\)JVI, \(\downarrow\)JR). For example, participants generating JI scores in excess of one standard deviation from the mean were found to form and express impressions that were typically harsher than 80% of the group and an average of 1.5–2 standard deviations above the mean. Judgment ratios below 1.0 provide evidence that the participant harbors a self-image that is more favorable than that held of others. Elevated JVI scores are thought to suggest a penchant to form extreme, often bipolar, evaluations of others.

1.3. Design and procedure

Participants were assigned to two separate interaction sessions with three classmates in each group. The sessions were scheduled at least one week apart, and participants were provided the following expectation:

You will be assigned to meet in a particular room at a particular time with three classmates. You will be instructed to simply interact freely with these classmates for a 25 min period which will hopefully provide an opportunity to get to know one another. At the conclusion of each interactive session, you will be provided a confidential space where you will be asked to provide your impressions, through various formats, of the personality attributes exhibited by each of your three partners. You also will be asked to make predictions about the manner in which each of your interaction partners rated you at the end of the preceding session.

Up to 36 participants began each session in a large classroom, followed by assignment of groups to individual rooms with color coded (by group) identification badges labeled A through D. Participants were informed that there were no restrictions on the conversation content and
that it was acceptable to share their name or other personal information with the group. At the conclusion of 25 min, the experimenter distributed a matching (by letter) color coded folder to each group member while ushering everyone back into the large classroom to sit apart and complete the rating forms provided in the assigned folder. Reusable laminated copies of the instructions and the 14-item FIIP rating scale assisted participant efforts to provide judgments, predictions and self ratings. A BASIC scoring program was developed to facilitate calculation of the three FIIP judgment variables.

Participants were invited to visit an assigned office to complete the MCMI-II at their own discretion which usually occurred during the first month of course enrollment. Testing was completed in privacy, and each participant coded the resulting protocol with a six-digit self-generated number (birth dates were suggested) to allow for anonymity but eventual matching of the scores with the FIIP data. Each class was monitored closely to ensure no redundancy in the use of these anonymous code numbers. No participant was given test feedback, and study objectives were described in only the most general terms.

2. Results

2.1. Descriptive statistics

Men (\(M = 2.54, \text{S.D.} = 0.71\)) provided significantly harsher partner ratings than women (\(M = 2.25, \text{S.D.} = 0.63\)), \(F(1, 183) = 6.86, \ P = 0.009\). Gender differences were not found for either JVI, \(F(1, 183) = 0.01, \ P = 0.92\), or JR, \(F(1, 183) = 0.23, \ P = 0.64\), scores which averaged 0.59(SD = 0.37) and 1.0(SD = 0.21) respectively. Men scored significantly higher on the antisocial, \(F(1, 183) = 4.22, \ P = 0.04\), and sadistic–aggressive, \(F(1, 183) = 5.08, \ P = 0.03\), scales, with women significantly higher in compulsive base rate scores, \(F(1, 183) = 4.84, \ P = 0.03\).

2.2. Correlation analyses

Table 1 presents the results of simple correlation analyses between the three FIIP judgment variables and the MCMI-II personality scales. JI scores were associated with higher self-defeating, schizoid, avoidant and schizotypal base rate scores among men, and antisocial, narcissistic, schizotypal, and passive–aggressive features among women. The JVI was linked to narcissistic, histrionic, antisocial, sadistic, schizoid, and paranoid scores among women. Antisocial base rate scores were positively correlated to the JVI for men. JR scores were inversely related to schizotypal and self-defeating scores among men and the histrionic, narcissistic, and antisocial scales among women. JR scores among women were positively correlated to the schizoid, sadistic, and avoidant scales.

Fisher’s \(z\) transformation procedure (Ferguson, 1981) was used to identify gender differences in correlation strength for 11 of the comparisons presented in Table 1. Significantly stronger relationships were found between the histrionic, narcissistic, and sadistic scales and the FIIP JR for women. Conversely, the self-defeating and schizotypal scales provided significantly stronger predictors of JR for men. JI was predicted more effectively by the narcissistic scale for women and self-defeating and schizoid scores for men.
2.3. DSM-IV personality disorder clusters

Table 2 provides the results of correlation analysis based upon MCMI-II personality scores that are clustered into the groupings suggested by the DSM-IV. Cluster B personality dimensions (antisocial, borderline, histrionic, narcissistic) were significantly related to all three FIIP judgment variables for women. Cluster A (paranoid, schizoid, schizotypal) provided a significant predictor of the JI, along with a trend for JR, for men.

2.2.1. Stepwise regression

The Judgment Ratio provides the purest FIIP measure of judgmental attitude and behavior since it incorporates self-image and predicted personal impressions as a control for the ratings of targets. The 13 MCMI-II personality scales were entered into separate stepwise regressions for men and women to predict JR scores. The best predictors of JR for women were the histrionic, compulsive, and narcissistic scales accounting for 12, 15, and 17% of the cumulative variance, respectively, in a three-step solution. The equation for men utilized the schizotypal, dependent, compulsive, self-defeating, and borderline scales for cumulative variance of 10, 14, 19, 26, and 30%.

3.3.2. Profile considerations

JR scores that are less than 1.0 indicate a tendency for participants to rate others more harshly than self. A series of 2 (gender) by 2 (JR < 1 versus JR = > 1) ANOVAs further illustrated gender
by JR interaction effects on MCMI-II histrionic, $F(1,181)=6.52$, $P<0.01$, and narcissistic, $F(1,181)=7.07$, $P<0.01$, base rate scores (see Figs. 1 and 2). Dependent personality base rate scores were significantly lower among both men and women with JRs below 1.0, $F(1,181)=3.78$, $P<0.05$.

King and Pate (2002) identified judgmental participants by JI scores in excess of one standard deviation from the mean (2.82 and 3.24 for women and men, respectively). Mean JI scores in this standardization sample were 2.52 (SD = 0.72) for men and 2.2 (SD = 0.62) for women. A series of 2 (gender) by 2(JI > 1 SD versus JI < M) ANOVAs culminated in findings that were also con-

---

**Table 2**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JI</td>
<td>JVI</td>
</tr>
<tr>
<td>A</td>
<td>0.09</td>
<td>-0.03</td>
</tr>
<tr>
<td>B</td>
<td>0.24**</td>
<td>-0.27***</td>
</tr>
<tr>
<td>C</td>
<td>0.02</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

JI, Judgement Index; JVI, Judgement Variability Index; JR, Judgment Ratio. Cluster A, paranoid, schizoid, and schizotypal; Cluster B, antisocial, borderline, histrionic, and narcissistic; Cluster C, Avoidant, dependent, and compulsive.

* Significant gender difference in correlation strength.

** $P<0.05$.

*** $P<0.01$.

---

**Fig. 1.** Mean scores among women with JRs below and above 1.0 was 75.3 (S.D. = 18.9, SE = 2.2, $N=74$) and 63.7 (S.D. = 21.5, SE = 2.7, $N=63$). Mean scores among men with JRs below and above 1.0 was 61.2 (S.D. = 19.1, SE = 3.7, $N=27$) and 66.5 (S.D. = 16.4, SE = 3.6, $N=21$).
consistent with the correlation results presented in Table 1. Gender by JI group interaction effects were found for the schizoid, \( F(1,119) = 3.89, P < 0.05 \), and narcissistic, \( F(1,119) = 4.09, P < 0.05 \), scales. Narcissistic elevations were more likely to be associated with harsher JIs among women. Schizoid elevations were more strongly associated with harsher JIs among men. Self-defeating, \( F(1,119) = 6.87, P < 0.01 \), and schizotypal, \( F(1,119) = 6.18, P < 0.01 \), scores were associated with harsher JIs among both men and women.

3. Discussion

The First Impression Interaction Procedure provides an opportunity to observe judgmental rating tendencies among individuals exposed to six unfamiliar partners in common unstructured social encounters. The FIIP Judgment Index provides an estimate of the harshness or leniency of the participant’s uninformed evaluations of others. The Judgment Ratio qualifies judgmental central tendency estimates with self-image ratings that provide a more sophisticated measure of the attitude of the evaluator. Judgmental tendencies are suggested primarily by JRs below 1.0 and JIs in excess of 1 SD from the mean. King and Pate (2002) found that men did tend to provide harsher JI ratings than women (about 0.5 SDs), but gender differences were not observed for JR or JVI scores. FIIP judgment scores did not seem to vary as a function of the gender balance within interaction groups, and random group assignment and combined data from two sessions is thought to provide internally consistent, temporally stable judgment scores.

MCMI-II profile elimination criteria corresponded with those recommended in the literature, and additional unreported analyses were conducted to assess whether the present results emerged largely as a function of exclusion decisions. The conclusions presented here were not altered significantly by the inclusion of these 41 discarded participants or the use of raw versus base rate.
MCMII scores. These exclusion criteria did, however, provide assurance that protocols distinguished by carelessness or social desirability would be excluded. The possibility that the reported FIIP and MCMII relationships emerged secondary to depression or anxiety was also considered. With only one exception, the MCMII dysthymia and anxiety measures were unrelated to the FIIP judgment variables. Dysthymia was significantly correlated with JR scores among men, \( r(46) = -0.30, p < 0.05 \). MCMII correlations with JR scores were marginally reduced (self-defeating and schizotypal relationships rendered insignificant) when variance associated with the dysthymia and anxiety scales was controlled, but the stepwise regression results remained identical. Most importantly, the gender interactions described in the results were unaltered by unreported covariance analyses which controlled for these variables.

3.1. Judgment ratio

The Judgment Ratio provides a unique index of judgmental attitude since it is derived from the ratio of ratings to self-image. Judgment ratios that are below 1.0 reflect a tendency to rate self, and expect others to rate self, more favorably than others. King and Pate (2002) demonstrated that JR scores were not simply a function of an objective analysis of observable partner attributes since participants with low ratios tended to provide ratings of others that were harsher than 80% of the group and an average of 1.5–2 standard deviations above the mean (e.g. minimal consensus validity with other group members).

Many of the MCMII personality dimensions provided significant predictors of the FIIP judgment variables, and gender differences were often found in the nature of these relationships. JR scores among women were associated with narcissistic, histrionic, schizoid, sadistic, and to a lesser extent, antisocial and avoidant features. Conversely, the schizotypal and self-defeating scales provided the best predictors of JR among the men. To this extent, judgmental attitude as measured by the FIIP was associated with DSM-IV cluster B (“dramatic, emotional, or erratic”) features among women and cluster A (“odd or eccentric”) qualities among men. These gender by JR interactions were particularly evident for the histrionic (see Fig. 1) and narcissistic (see Fig. 2) scales. Mean histrionic and narcissistic scores for men did not differ significantly between participants generating JRs above or below 1.0, but scores were significantly higher among women with JRs below 1.0. Women with JR scores below 0.8 could be predicted with some confidence (14/20 or 70%) to have a MCMII histrionic elevation (BR > 74). Five out of six (83%) of the women with histrionic base rate scores above 94 scored below 1.0 on the JR. These associations were less pronounced for the narcissistic scale where the differences between men and women were primarily a function of scores among nonjudgmental participants (JR⇒1).

3.2. Judgement index

The Judgment Index can be conceptualized as a simple measure of judgmental behavior that may or may not reflect negative attitudes toward the target(s). As Cronbach (1955) observed, the meaning of dimensional ratings can, and typically do, vary as a function of the personal standards applied by the judge. The assignment of academic course grades offers a good example. One professor rarely assigns “A” grades and considers the “B” student to be exceptional. The same “B” grade given by another college professor might reflect a different judgment that the
student’s performance is relatively weak compared with the remainder of the class. Students receiving grades below personal standards and those assigned to classmates may nevertheless characterize the ratings as judgmental even in cases where the evaluator was not attempting to be critical.

JI scores among women were associated with antisocial, narcissistic, schizotypal, and passive–aggressive features. Once again, men tended to form close associations with almost entirely different scales (self-defeating, schizoid, avoidant, and schizotypal). Statistically significant gender differences in correlation strength were found for the narcissistic, schizoid, and self-defeating scales. Rating behavior appeared to be differentially coarsened by personality traits of social alienation in men and grandiosity in women.

3.3. Judgment variability index

The Judgment Variability Index is calculated as the standard deviation of JI scores generated for each of the six partners. The JVI represents the degree to which participants discriminated in their application of JI ratings and provide an estimate of the degree to which situational factors such as partner attributes, mood state changes, and other factors influence first impressions. Low JVI scores suggest inclinations to judge others similarly despite the many differences which are certain to occur in partner behavior, conversational content, and other circumstances that will vary from session to session. JVI elevations assure that one, or usually more, of the six partners were rated harshly relative to the remainder of the group. Indeed, JVI was found by King and Pate (2002) to be significantly correlated to JI ($r = 0.46$) and JR ($r = -0.41$) in directions that support its employment as a unique form of judgmental attitude and behavior.

JVI scores among women were associated with narcissistic, histrionic, antisocial, sadistic, paranoid, and schizoid features. JVI scores were predicted only by the antisocial scale among men, but gender differences in correlation strength were not found for any of the personality dimensions. Higher variability in FIIP ratings tends to be associated with lower JR and higher JI scores, as well as higher scores on a host of MCMI-II personality dimensions.

4. Summary

MCMI-II interpretive manuals (Choca et al., 1992; Millon, 1987) and the DSM-IV Axis II diagnostic criteria provide a basis for predicting associations between judgmental attitudes and behavior and a number of personality trait dimensions (e.g. antisocial, passive–aggressive, paranoid, narcissistic, borderline, compulsive, etc.). Personality effects on judgmental behavior appeared to be manifested differently among men and women. The MCMI-II histrionic scale proved to be a surprisingly strong predictor of the JR among women. These results suggested that personality variables may mediate impression formation in a different way depending on the gender of the observer. Personality features such as social alienation and insecurity provided the best predictors of judgmental tendencies among men, while harsh FIIP judgment ratings were more common among self-confident women who presumably sought attention and affirmation from others. Attempts should now be made to replicate these findings within clinical samples.
References


