Social Network Analysis of Self and Peer Perceptions of Personality Pathology

Allan Douglas Clifton
Charlottesville, VA

B.A., Haverford College, 1995
M.A., University of Virginia, 1999

A Dissertation Presented to the Graduate
Faculty of the University of Virginia
In Candidacy for the Degree of
Doctor of Philosophy

Department of Psychology
University of Virginia

January 2005

Thomas F. Oltmanns

Eric Turkheimer

Thomas F. Oltmanns

Timothy D. Wilson

Timothy Wilson

Thomas M. Guterbock

Thomas Guterbock

Abstract

Personality disorders (PDs) are most often evaluated on the basis of self-report, despite involving the way that one's behavior affects others. Nearly all studies of peer perceptions of PDs have relied on self-selected informants, which may result in low reliability and overly positive biases. Although obtaining information from large groups of unselected peers is preferable, it introduces complicating effects of group dynamics. In addition, in a large group of peers, not all raters will make equally valid ratings of all targets.

The present study utilizes social network analysis to investigate ways of improving reliability and validity in peer ratings. Participants were 21 groups of peer raters from a military population (*N*=809) who acted as both targets and raters in a round-robin design. Using the Peer Inventory for Personality Disorder, individuals identified other participants who exhibited traits of DSM-IV personality disorders. Participants also completed self-report versions of the same instrument. Mixed linear models were used to estimate variance in peer ratings due to network, rater, target, rater-target interaction, and self-report.

Adjacency matrices were constructed based on participants' self-report of how well acquainted they were with one another. Social network analysis was then applied to find network characteristics of participants, and to identify a variety of cohesive subgroups within networks.

Network characteristics were associated with both self- and peer-reported personality disorder traits. Consistent with DSM-IV descriptors, measures of centrality

and degree connectivity were positively associated with narcissistic and histrionic PDs, and negatively associated with avoidant, schizoid, and schizotypal PDs.

Peer ratings made within cohesive subgroups were larger, had higher self-peer agreement, and were more reliable, than did those made by raters who did not share a mutual subgroup with the target. Partitioning networks into two subgroups achieved improvements as large as, and more consistent than, identifying small tight-knit cohesive subgroups.

Social network analysis is posited as a means of incorporating aspects of Kenny's (1994) Weighted-Average Model in a cruder, but more parsimonious, way. It is recommended that researchers investigating peer perceptions of normal and abnormal personality consider partitioning large groups into two cohesive subgroups, to maximize reliability and validity of ratings.

Table of Contents

List of Tables

List of Figures

Acknowledgments

Data for this project were collected as part of a larger study, “Peer Assessment of Personality Traits and Pathology,” funded by the National Institutes of Mental Health grant number 1R01MH51187-06. I would like to thank the principal investigator and my major advisor, Eric Turkheimer, Ph.D., and co-investigator Thomas F. Oltmanns, Ph.D., for their direction and support throughout my graduate studies. I would also like to express my appreciation to the other members of my dissertation committee, Timothy Wilson, Ph.D., and Thomas Guterbock, Ph.D., for lending their time, expertise, and tough questions to help shape the finished product.

My wife Elizabeth was a vital contributor to this dissertation, providing love, encouragement, and understanding throughout the long process. I am also grateful to my family (particularly my parents, Wayne and Nancy Clifton), my friends, and the other members of my cohesive subgroups for their support along the way.

Introduction

Personality disorders (PDs) are maladaptive patterns of behavior, cognition, or interpersonal functioning which lead to significant impairment or distress (American Psychiatric Association, 1994). PDs are quite common, with an estimated 10% to 14% point prevalence of personality disorders in nonclinical populations (Lenzenweger, et al. 1997; Weissman, 1993). In both research and clinical settings, PDs are most often diagnosed on the basis of self-report, obtained through written inventories or clinical interview. However, the nature of PDs inherently involves the way in which one's maladaptive personality affects others (Westen, 1997), which individuals with PDs may have an especially difficult time observing or reporting (John and Robins, 1994; Oltmanns, Turkheimer, & Strauss, 1998). Further, the criteria used to rate personality disorders tend to be highly evaluative, which may lead to defensiveness and cognitive distortions in self-report (John and Robins, 1993; Kenny and Kashy, 1994). Relying solely on self-report methods of assessment, therefore, may produce a limited view of PD traits.

Obtaining information from peers provides an alternate view of the interpersonal aspects of personality pathology. Peer perceptions of pathological personality traits are usually obtained from a knowledgeable informant, who describes the personality of the participant via questionnaire or structured interview (McCrae & Costa, 1987; Zimmerman, Pfohl, Stangl, & Corenthal, 1986). This methodology has two major limitations. First, it obtains information from only one or two informants, which necessarily limits the reliability of the data. Second, informants selected by the participant may suffer from what has been described as the "letter of recommendation" problem (Klonsky, Oltmanns, & Turkheimer, 2002). That is, the close friends, spouses, or relatives who are chosen as informants may tend to describe participants in a positive light. Unselected peers who interact with the individual on a regular basis, such as co-workers or classmates, are likely to be more objective in their judgments.

Obtaining reports from multiple peers, as opposed to individual informants, is not uncommon in the assessment of normal personality (e.g., Mount, Barrick, & Strauss, 1994; Strauss, Barrick, & Connerley, 2001; Brutus, Fleenor, & McCauley, 2001). However, ours is the only project to date to gather information about maladaptive personality traits from a large group of unselected peers (Klonsky, et al., 2002). This methodology presents a more complete picture of the interpersonal consequences of pathological personality than informant sources.

A possible drawback to the use of unselected peers is the complicating effect of group dynamics. Ratings in large group studies may be affected by a variety of interpersonal variables, including rater-rater acquaintance (e.g., Kenny and Kashy, 1994; Kenny, Albright, Malloy, and Kashy, 1994), rater-target acquaintance (e.g., Park and Judd, 1989), degree of overlap in observations by raters (Kenny, 1994), in-group and out-group effects (e.g., Tajfel, 1978), differing meaning systems among raters (Kenny, 1994), differing average ratings by different raters (e.g., Albright, Kenny, and Malloy, 1988), and a host of other factors.

In studies of individual informants, factors such as these may be more easily taken into account, as informants can be interviewed in depth regarding their association with the target. In laboratory studies of person perception at low levels of acquaintance, many of these factors may also be calculated and monitored using specialized techniques such as the Social Relations Model (Kenny and LaVoie, 1984). For example, given perceivers previously unacquainted with the target, behavioral perception overlap may be manipulated by allowing perceivers to observe particular subsets of the target's behavior. Similarly, given perceivers previously unacquainted with one another, communication among raters may be manipulated by circumscribing the amount or type of communications they engage in.

These factors become more difficult to interpret in larger, ecological group studies of personality perception. When participants have previously known one another for significant amounts of time, they may have interacted in any number of settings and situations, creating an unknown amount of similarity among perceivers' interactions with the target. The amount of communication among two perceivers regarding any given target is similarly difficult to ascertain. This is not to say that such information is impossible to obtain, simply that obtaining it requires more effort than most researchers can afford. To evaluate the effects of communication in large groups of acquaintances requires that each judge be asked about their communication with each other judge regarding each target. In a round-robin design of 40 participants, this would involve (40x39x38) = 59,280 additional pieces of information, all of which are retrospective estimates by participants, with the additional error this entails (e.g., Bernard, Killworth, Kronenfeld, and Sailer, 1984). Assessing perception overlap, agreement about stereotypes, and shared meaning systems in perceivers is equally difficult and time-consuming.

In group studies of personality perception, treating all judges as equally good raters, or accounting only for acquaintance, may overlook valuable sources of information (Kenny, 1994). In order to take into account some of the mediating effects of the interpersonal environment, information obtained from social network analysis of the data may act as a surrogate for more specific factors such as overlap and communication. Kanfer and Tanaka (1993) advocated research on the effects of social networks in personality assessment:

> A systematic program of research applying methods from social network analysis to help personality researchers unravel the connections between personality assessment and social context promises to yield a variety of benefits. Identification of particular positions that afford greatest access to social information in specific contexts will allow researchers to select the most accurate judges for personality assessments. Moreover, such a research program can address fundamental questions about the social nature of personality constructs, including the perception of self and others. (pp. 735-736)

Although social network analysis (SNA) has been widely used in a variety of related disciplines, it has rarely been applied to interpersonal perception (Kanfer and Tanaka, 1993). The present study uses SNA to investigate the systematic differences which may arise in ratings of personality pathology by groups of peers. The ability to predict these differences would contribute to improved assessment of personality disorders, and to greater understanding of personality pathology in general.

Self-Peer Agreement in Personality Disorders

Prior studies of informant assessments of personality disorders have demonstrated only modest agreement between self and peers. A recent review identified 30 published studies of self and informant reports of personality disorders (Klonsky, Oltmanns, and Turkheimer, 2002). The review concluded that self-informant correspondence was "modest at best" (Klonsky et al., 2002, p. 308) with a median correlation of r=.36 in studies of DSM personality disorders. The review found greater self-peer agreement for the Cluster B personality disorders (median r=.45) than for either Cluster A or C (median r=.35 for each). Studies of non-DSM personality pathology, such as trait scales for the Schedule for Nonadaptive and Adaptive Personality (Clark, 1993), also demonstrated slightly higher self-informant correspondence, with a median r of .47.

In a study of personality disorder in depressed inpatients, (Zimmerman, Pfohl, Coryell, Stangl, & Corenthal, 1988), 66 patients and 66 informants (friends, spouses, or first-degree relatives) were interviewed using the Structured Interview for DSM Personality (SIDP). Diagnostic concordance between patient and informant was poor for both individual personality disorders (kappa range = -.06 to .28) and for the presence of any personality disorder (kappa = .13). Comparing the dimensional scores, rather than categorical diagnoses, increased the agreement somewhat, both for individual PD's (correlations ranged from .15 to .61), and for the total number of traits endorsed (r=.53). Overall, informants reported significantly greater numbers of personality disorder traits than did patients. The researchers suggest that this discrepancy may be due to patients denying negative traits, although they note that criteria for Antisocial PD, which are quite socially undesirable, were endorsed with equal frequency by patients and informants.

Dowson (1992) conducted a questionnaire-based study of personality disorders, with a major focus on narcissistic PD (NPD). 60 clinical participants and 60 informants (generally spouses or first-degree relatives) were given the appropriate self or informant form of the revised Personality Diagnostic Questionnaire (PDQ-R). There was a moderate but significant correlation between self and other for both the total number of NPD items endorsed (r = .42) and the overall number of PD traits endorsed (r=.48). Dowson also compared the self-other correspondence for individual NPD items. Of the nine NPD criteria, only one, "Has a sense of entitlement" showed significant agreement between self and informant (kappa = .62).

Riso, et al. (1994) recruited 105 outpatients with either a history of mood disorder or personality disorder without mood disorder. Participants were interviewed with the Personality Disorder Examination (PDE). Each participant also provided one informant who was interviewed about the participant, also using the PDE. The diagnostic agreement for the presence of any personality disorder was a kappa of 0.06. Lowering the threshold for informants' diagnoses to below the DSM-III threshold increased this kappa to 0.20. Comparing the dimensional scores yielded somewhat higher correspondence, with a mean intraclass correlation coefficient of 0.35. The ICC for individual diagnoses varied quite a bit, however, where antisocial PD had the highest correspondence (ICC=0.55) and obsessive-compulsive PD had the least (ICC=0.19). Comparing the behavioral items of the PDE with the non-behavioral items showed a slightly higher correlation for the behavioral items (ICC=0.51 vs. 0.35). Comparing items sorted into low versus high social undesirability found no difference between the two (ICC=0.44 vs. 0.40).

Ferro and Klein (1997) administered a battery of personality disorder interviews and questionnaires to 224 probands, acting as informants, and 1,070 first-degree relatives. Proband informants reported at least one PD in 34% of the relatives, whereas only 22% of relatives self reported PD's, with the greatest discrepancy lying in narcissistic PD. Further, the kappa for individual PD diagnoses ranged from -.01 to .28, with a median value of .10. The kappa for any PD diagnosis at all was .16. Using a dimensional measure of the number of PD criteria met increased agreement somewhat, with correlations ranging from .14 to .40 (median = .18). Overall, however, this study again suggests that the concordance between self and informant report for personality disorders is relatively modest.

To date, the only published large group study comparing self and peer perspectives of personality pathology was by Oltmanns, Turkheimer, and Strauss (1998). The authors administered both self- and peer-report measures of personality disorders to 3 nonclinical samples of women (*N*s = 41, 58, and 63). Three PDs (narcissistic, NPD; dependent, DPD; and obsessive-compulsive, OCPD) were assessed, using modified versions of the SCID-II Questionnaire for self-report and the Peer Inventory for Personality Disorder (described below) for peer report. In the peer nomination portion of the study, each participant nominated three members of her group for each trait. Combining all participants across samples showed that there was low correlation between self-report and peer measure within each diagnostic category (NPD = .13, DPD = .12, OCPD = .30). One limitation of the study by Oltmanns and colleagues (1998) is the aggregation of peer data. That is, the peer reports were combined before self-peer correlations were calculated, which obscures any differential rater effects. The present study instead examines individual ratings of each target by each rater, allowing effects of social structure to be computed.

Social Network Analysis

Traditional assessments of personality perception have generally focused on dyads, in which one or two informants rate the personality of a target person. Originally an outgrowth of the sociometry techniques of Moreno (1934), social network analysis (SNA) looks at all of the targets and raters within an entire social system, focusing on the patterns of relationships (Kanfer & Tanaka, 1993). SNA has been used extensively in a wide range of disciplines including sociology, anthropology, economics, marketing, and engineering (Wasserman & Faust, 1994). However, despite promising possibilities in the area of personality assessment (Kanfer & Tanaka, 1993), SNA has rarely been applied to this task (Funder & West, 1993). This section will outline some basic principles of SNA, before proceeding to describe the application of SNA to informant studies in general and the present study in particular.

SNA analyzes the patterns of relationships within an entire bounded social network. Such a network might be defined as the employees in a workplace, the members of a village, students in a sorority, or any other group of individuals around which a meaningful boundary can be drawn. The boundaries of a sample are not always easy to determine (Scott, 2000), but it is necessary that there be meaningful relations among the members of the sample (Kanfer & Tanaka, 1993).

Given this bounded network, SNA can be used to interpret an individual's relationship to the entire group, rather than just his or her relationship to another individual. For example, within a network, subgroups of individuals with strong interrelations to one another can often be identified (a very informal definition of a "clique"). An individual who is connected to (i.e., has relationships with) only the people within a subgroup might have very different personality characteristics than an individual with connections to many different subgroups throughout the network. In this example, SNA could be used to identify the different patterns of relationships that these two individuals tend toward.

Social network terminology

Although social network analysis has been widely used in related social science fields, many of the terminologies used in SNA may be unfamiliar to those in the field of clinical psychology. I will therefore summarize some of the main concepts before proceeding to discuss their application to personality perception.

Sociomatrix

A sociomatrix, also called an "adjacency matrix," is a square matrix representing the relationships between the individuals in a network. Participants in a social networking study are asked to describe their relationships with each other member of the network, via ratings, nominations, or rankings. Participants might be asked whether they know each person, how much they like each person, how much they trust each person, how likely they would be to ask each person for advice, or virtually any other type of social relation. Each participant's response is then entered into the sociomatrix. A sociomatrix is by definition a round-robin design, such that all individuals in the network are both respondents and potential targets. A sample dichotomous non-directed sociomatrix appears in Table 1.

Table 1
Sample dichotomous non-directed sociomatrix.

	Participant Number															
Participant Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	--	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
2		--	1	1	1	0	0	0	0	0	0	0	0	0	0	0
3			--	1	1	1	0	0	0	0	0	0	0	0	0	0
4				--	1	1	0	0	0	0	0	0	0	0	0	0
5					--	1	0	0	0	0	1	0	0	0	0	0
6						--	1	1	0	0	0	0	0	0	0	0
7							--	0	1	0	0	0	0	0	0	0
8								--	0	1	0	0	0	0	0	0
9									--	1	0	0	0	0	0	0
10										--	0	0	0	0	0	0
11											--	1	0	1	0	0
12												--	1	0	1	1
13													--	1	0	0
14														--	1	1
15															--	0
16																--

In a network of k individuals, the sociomatrix describing their relationships is a k by k square matrix, with the k participants arrayed in the same order across the rows and columns. By convention, the rows represent the respondents, and the columns the targets. Therefore, person i's rating of his or her relationship with person j would be represented in cell (i,j) of the matrix (Scott, 2000). The cells along the main diagonal of the matrix represent the individual's relationship with him or herself, and are generally (though not always) undefined (Kanfer & Tanaka, 1993).

The actual content of the cells varies depending on the type of relational data collected. The two major distinctions in data types are whether the data are binary or valued, and whether they are directed or non-directed (Scott, 2000).

Valued/binary

Relational data can be either valued or binary. Binary data simply indicates whether or not a relationship exists, represented as 0 or 1 in the sociomatrix. Valued data, on the other hand, describes the relative strength, intensity, or frequency of a relationship (Wasserman & Faust, 1994). For valued data, participants rate or rank order each relationship, which is represented numerically in the sociomatrix. Although many SNA procedures can be performed on valued data, some require binary data. In such cases, valued data can be dichotomized by choosing a cut-off point at which relationships are considered present or absent (Scott, 2000).

Directed/non-directed

The other major distinction in relational data is whether the data are directed or non-directed. Relationships among individuals are not always reciprocal. Person *i* may consider person *j* a close friend, whereas person *j* considers person *i* merely an acquaintance.

In SNA, the lack of reciprocity in relational data is represented by directed data, allowing person *i*'s rating of person *j* to differ from *j*'s rating of *i*. A directed sociomatrix may be asymmetrical, such that cell (*i*,*j*) does not necessarily equal cell (*j*,*i*).

In non-directed relational data, the relationship between two individuals is considered to be reciprocal. A non-directed sociomatrix is symmetrical around the main diagonal, such that cell (*i*,*j*) is equal to cell (*j*,*i*). Some SNA procedures can be performed only on non-directed data. In these cases, directed data may be transformed to non-directed data in a variety of ways (Wasserman & Faust, 1994). For binary data, asymmetrical data may be made symmetrical ("symmetrized") by coding a relationship if either partner reports one, or by coding a relationship only if both partners report one. For polytomous data, the higher of the two partners' ratings may be used, the lower of the two may be used, or the two may be averaged (Borgatti, Everett, and Freeman, 1999). Differing methods of symmetrizing the data may be used depending upon the stringency desired for the relationship.

Graphs

Another way of depicting social network data is via a graph. In a graph, individuals in a network are represented by points, or "nodes," in space. Relationships among individuals are drawn as lines between two nodes. The direction, distance, and length of these lines are arbitrary; the patterns of connections are important, rather than the physical locations (Scott, 2000). A simple graph, corresponding to the data in Table 1, is depicted in Figure 1.

Like an adjacency matrix, a graph can be either directed (a "digraph") or non-directed, and either binary or valued. In a non-directed graph, connections between nodes are depicted as simple lines. A directed graph indicates the direction of connections using arrows on the lines. If person *i* indicates person *j* as a friend, but person *j* does not reciprocate, they would be connected by an arrowed line directed towards node *j*. If both *i* and *j* indicate the other as a friend, their nodes would be connected by a double-headed arrow.

Binary connections are indicated by either the presence or absence of a line. Valued connections are depicted by a line with an accompanying number, indicating the strength of the connection.

Nodes (individuals) which are connected to one another are said to be “adjacent.” A series of lines connecting two nodes, in which each node along the lines is distinct, is called a “path.” (Scott, 2000). The shortest path between two nodes is called its “geodesic,” and the length of the geodesic is the “distance” between the two nodes. Two nodes which are adjacent are separated by a distance of 1, because they are directly connected. If two nodes are not directly connected, but have one node between them, the length of their geodesic, and therefore the distance between them, is 2.

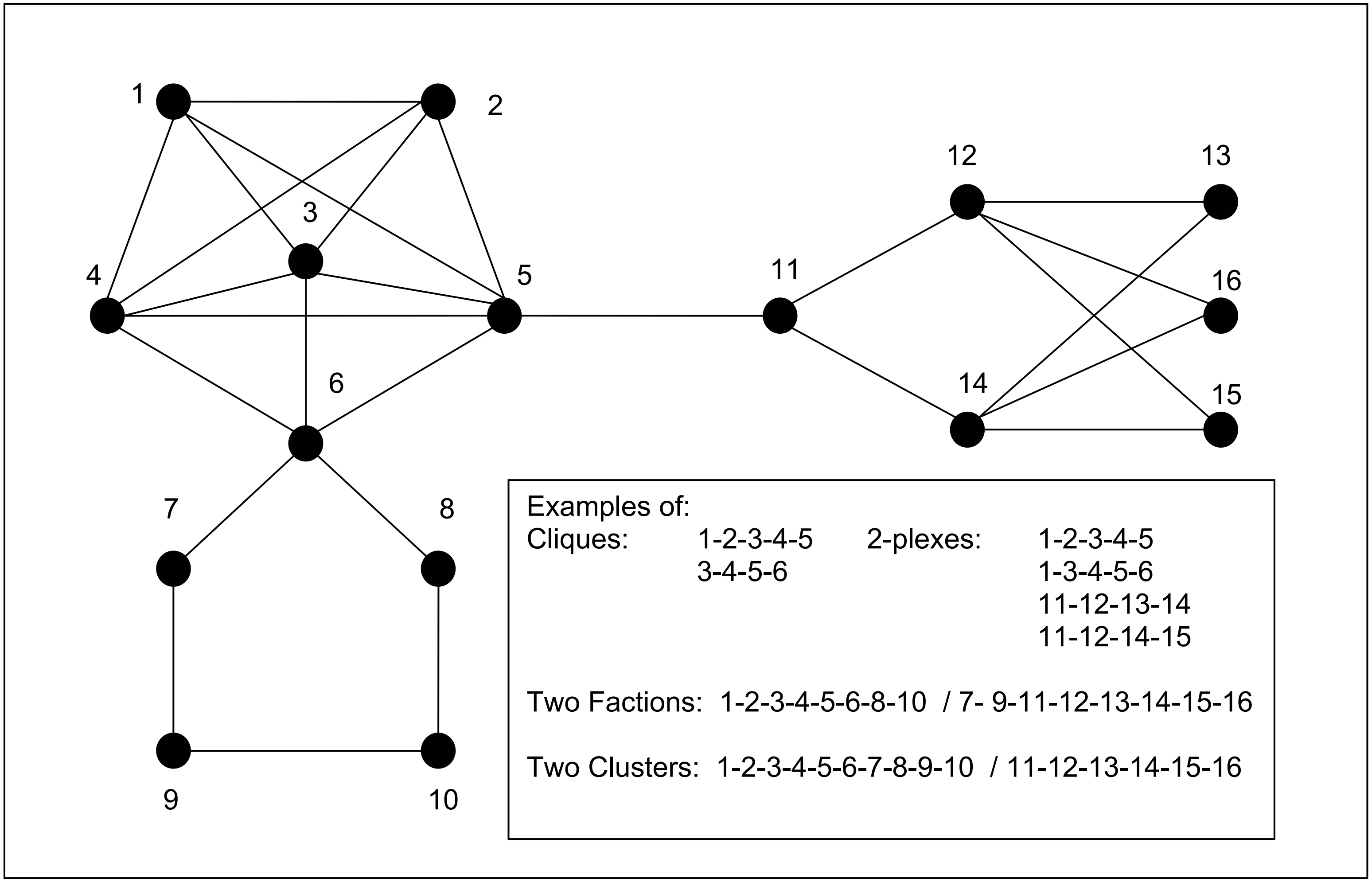

Figure 1. Sample Binary Non-Directed Graph Corresponding to Table 1.

Indegree/Outdegree

One of the most fundamental concepts in SNA is that of degree of connection (Kanfer & Tanaka, 1993). A node's degree is the number of other nodes adjacent to it. That is, an individual's degree is the number of other individuals directly connected to him or her (Wasserman & Faust, 1994).

In directional relations, degree can be broken down based on whether the connections are those that individual reported to others, or whether they are connections that others have reported to the individual. The former is called "outdegree," referring to the number of connections originating from the node. The latter is called "indegree" and represents the number of connections directed toward the node.

Degree, indegree, and outdegree can also be calculated using the adjacency matrix. In a binary non-directed sociomatrix, the degree of a node is equal to either the row sum or the column sum (which are by definition equal). More formally, given an adjacency matrix *X*, containing *g* members, the formula (Wasserman & Faust, 1994, p. 163) for the degree (*d*) of the node n_i is:

$$d(n_i) = \sum_{j=1}^{g} x_{ij} = \sum_{i=1}^{g} x_{ij}$$

In a binary directed sociomatrix, the indegree of a node is equal to its row sum, and the outdegree is equal to its column sum. More formally, the formulae (Wasserman & Faust, 1994, p. 164) for the indegree (d_I) and outdegree (d_O) are:

Indegree: $d_I(n_i) = \sum_{j=1}^{g} x_{ij}$ Outdegree: $d_O(n_i) = \sum_{j=1}^{g} x_{ji}$

Density

One of the most widely-used concepts in SNA is that of density (Scott, 2000).Whereas indegree and outdegree represent the connections of individual members of a network, density describes the number of connections across the entire network. In a network, the number of possible connections among members is constrained by the total number of individuals in the network. In a network with g members, the maximum number of binary connections in the network is $g(g\text{-}1)/2$ (Scott, 2000).

Density represents the proportion of actual connections within the network to the maximum number of connections possible. Density ranges from 0 (if no lines are present) to 1 (if all possible lines are present). Given the number of connections (*L*), density (represented by Δ) is equal to (Wasserman & Faust, 1994, p. 101):

$$\Delta = \frac{L}{g(g-1)/2} = \frac{2L}{g(g-1)}$$

In directed graphs, the maximum number of possible connections is doubled, because the direction of the connections are taken into account. In a directed graph, therefore, the maximum number of connections is g(g-1), and the formula (Scott, 2000, p. 71) for density is:

$$\Delta = \frac{L}{g(g-1)}$$

Centrality

Another important measure in SNA is centrality, which refers to the intuitive notion that some members of a network are central to the structure, while others are more on the fringe of the network. Centrality is a difficult concept to formally define, as reflected by the numerous competing models which have been proposed (e.g., Freeman, 1979; Nieminen, 1974; Bonacich, 1987; Sabidussi, 1966). The present work will utilize the "betweenness" model of centrality developed by Freeman (1977; 1979). This model is the most frequently used measure of centrality (Wasserman & Faust, 1994), and tends to produce both the most accurate results and the greatest variance in individual centrality scores (Freeman, 1979).

Freeman's (1977; 1979) model of centrality focuses on the "betweenness" of individuals. In a graph of a social network, not all nodes (i.e., individuals) are directly connected to one another. However, some may be connected indirectly, because both are connected to the same mutual node. In other words, person *j* may not know person *k* directly, but they may have a mutual friend in common. If the shortest path from *j* to *k* goes through node *i*, then *i* is said to be between *j* and *k*. Freeman (1979) argued that if a node lies between many other nodes, its greater "betweenness" makes it more central to the network than a node that is not between any other nodes. In Figure 1, node 5 has the greatest amount of betweenness centrality in the network, whereas nodes 1 and 2 have the least. Individuals with high betweenness may act as "power brokers" or "gatekeepers," controlling the relationships among other individuals in the network (Freeman, 1980). The concept of betweenness is also essential to the personality correlates of structural holes (Burt, Jannotta, & Mahoney, 1998), described below.

The betweenness measure can be used to calculate centrality in directed networks (Gould, 1987). Betweenness is based on finding the shortest possible path which connects two nodes, called the "geodesic." In concept, betweenness represents the probability that a given node lies on a geodesic connecting two other nodes (Wasserman & Faust, 1994).

Formally, the number of shortest-path geodesics connecting *j* and *k* is represented by g_{jk}. The number of shortest-path geodesics connecting *j* and *k*, of which *i* is a part, is represented by $g_{jk}(n_i)$. The probability that *i* lies on any given geodesic between *j* and *k* is therefore estimated as: $g_{jk}(n_i)/g_{jk}$ (Freeman, 1979).

The betweenness (C_B) for individual i (n_i) is calculated as the sum of the probabilities that *i* lies on the geodesic between any pair of nodes (Wasserman & Faust, 1994):

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i) / g_{jk}$$

$C_B(n_i)$ is often standardized by dividing by the maximum possible number of pairs of actors, not including n_i. This value, $C'_B(n_i)$, ranges from 0 to 1, and allows comparisons across networks (Wasserman & Faust, 1994). $C'_B(n_i)$ is calculated (Wasserman & Faust, 1994, p. 201) as:

$$C'_B(n_i) = C_B(n_i) / [(g-1)(g-2)/4]$$

Subgroups: Cliques and k-plexes

Individuals in large social networks often form into smaller social groups of closer friends (Johnsen, 1986). One of the most longstanding and widely used applications of SNA is the endeavor to identify these naturally occurring subgroups

(Scott, 2000). However, precisely how such a group should be operationalized has been widely debated. Wasserman and Faust (1994, p. 252) noted four separate characteristics that have been used to operationalize cohesive subgroups. They are:

- “The mutuality of ties
- The closeness or reachability of subgroup members
- The frequency of ties among members
- The relative frequency of ties among subgroup members compared to non-members”

Each of these characteristics has been used to construct definitions of different types of social subgroups. I will briefly review examples of each, with specific attention paid to the techniques utilized by the present study.

The first of the properties cited by Wasserman and Faust (1994), mutuality of ties, is embodied by the concept of the clique. In social network analysis, a clique is defined as a social subset of at least three members, in which every individual is adjacent to every other individual (Doreian, 1979). Individuals may be members of more than one clique, but two cliques cannot overlap completely, as the larger would subsume the smaller (Wasserman & Faust, 1994). Because they require that each member be adjacent to each other member, cliques are rather unstable. Because the requirements are restrictive, simply changing one member of a clique can alter the structure of the entire subgroup. Cliques have therefore been largely supplanted by less restrictive methods of assessing cohesive subgroups (Wasserman & Faust, 1994).

The second subgroup property cited by Wasserman and Faust (1994) is closeness or reachability among subgroup members. Reachability is based on the premise that even if individuals are not directly connected to one another, information may still pass between them through intermediaries or mutual acquaintances (Wasserman & Faust, 1994). Techniques which use this property to operationalize cohesive subgroups include n-cliques and n-clans (Wasserman & Faust, 1994). Both are less restrictive extensions of the clique concept, in which the requirement that all members be directly adjacent to one another is relaxed. N-clique and n-clans are subgroups in which the maximum distance between any two members is n. Therefore, in a 2-clique, all members must either be directly adjacent to one another, or connected through a mutual friend, making them separated by a maximum distance of 2 (Doreian, 1979). The n-clan is slightly more restrictive, in that it requires that all members who are not directly connected must be connected through another member of the n-clan (Mokken, 1979). That is, persons A and B may be members of a 2-clan by the their connection through person C, only if person C is a member of the clan. Both n-cliques and n-clans have been used in the study of information flow, but are less useful when studying direct associations among individuals (Wasserman and Faust, 1994). These measures were therefore not incorporated into the present study.

The third method property that Wasserman and Faust (1994) cite as being used to define cohesive subgroups is the frequency of ties among members. The intuitive understanding of a cohesive subgroup suggests that members of a subgroup will know all or most of the other members of the group. For instance, the definition of a traditional clique requires that all members are directly adjacent to all other members. That is, in a clique with g members, every member must be directly connected to (g-1) members (every member except him or herself). The concept of the k-plex, proposed by Seidman and Foster (1978), relaxes this property of the clique, and requires instead that every member be directly connected to (g-k) other members. In other words, a 2-plex requires that every member be adjacent to all but one of the other members (plus the member him or herself). Unlike n-cliques and n-clans, these connections must be direct; connections within the k-plex cannot go through intermediaries. The present study utilizes 2-plexes as a measure of cohesive subgroup.

The final method of operationalizing cohesive subgroups discussed by Wasserman and Faust (1994) is comparing within-group ties to between-group ties. This method assumes that connections within a cohesive subgroup will be stronger or more frequent among members of a subgroup than among a member of a subgroup and a non-member. One application of this methodology is the lambda set, introduced by Borgatti, Everett, and Shirey (1990). Lambda sets are based on the idea that cohesive subgroups should be robust in their connections, and difficult to disconnect. That is, if a pathway between two individuals is removed, alternate pathways between the individuals should remain. Unlike the cohesive subgroups described above, lambda sets are not dependent on adjacency or limited by geodesic distance (Wasserman & Faust, 1994). Lambda sets are therefore distinct from the intuitive notion of a clique, in that members of a lambda set may not know one another, and in fact may be quite distant from one another, so long as there are multiple pathways connecting them (Borgatti, et al., 1990). Because lambda sets do not require direct contact between individuals, they were not included in the present study.

Cluster Analysis

Cluster analysis is a generic term to describe a varied group of statistical procedures. All of the numerous varieties of cluster analysis start with data about some number of items, and attempt to reorganize them into relatively homogenous groups (Aldenderfer & Blashfield, 1984). The means by which this is accomplished varies by the type of cluster analysis used. Even the precise meaning of "cluster" varies by the type of cluster analysis, as clusters may be overlapping (in which multiple clusters may include the same members), hierarchical (which precludes overlapping membership, although one cluster may wholly contain another), or disjoint (in which objects may be members of one and only one cluster) (SAS Institute, 1999).

Various methods of cluster analysis have been used to find higher density subgroups within social network data (Scott, 2000). The present study uses a nonparametric clustering method via the MODECLUS procedure in the SAS statistical package. This procedure is a disjoint method, meaning that for any number of clusters in a network, each individual will be assigned to one and only one cluster. MODECLUS converts the adjacency matrix into a distance matrix, such that a lower knowing score is interpreted as a greater distance between individuals. The procedure then uses an agglomerative procedure (Scott, 2000) in which small regions ("kernels") of local maximum density are joined with other kernels to form larger regions of local maximum density (SAS Institute, 1999). The density function of each kernel is estimated by

$$\hat{f}_i = \frac{n_i}{nv_i}$$

where n_i is the number of neighbors of x_i, n is the number of individuals in the network, and v_i is the volume of the kernel (SAS Institute, 1999). The size of the kernels which are subsequently joined together to form clusters is based on a user-specified smoothing parameter, which restricts the radius and/or minimum number of individuals in a kernel.

Cluster analysis techniques are useful tools for social network analysis, particularly because they are available in most statistical software packages, without the need to obtain specialized SNA software. In addition, nonparametric cluster analysis methods like MODECLUS are in many ways superior to other clustering techniques, particularly in identifying clusters of unequal sizes or irregular shapes (SAS Institute, 1999). However, one of the limitations of cluster analysis is its reliance on local maxima for density estimation. If, further into the agglomerative process, a better solution using different kernels were to arise, the procedure cannot take a step backwards and change the an earlier step to improve the ultimate solution.

Factions and TABU search

A final method of unearthing cohesive subgroups is the Factions procedure within the UCINET 6 analytical program (Borgatti, Everett, and Freeman, 2002). Factions is based on a recently developed search algorithm for large combinatorial problems, called "Tabu". The Tabu algorithm is a search heuristic that seeks to optimize some function, by beginning with a random or specified solution and "moving" in a direction that improves the function. Moves continue until no further moves improve the function (Salhi, 1998). For instance, Tabu and similar heuristics have been used to solve traveling salesman-type problems (e.g., Lawler, Lenstra, Kan, & Shmoys, 1990; Malek, Guruswamy, Owens, & Pandya, 1989), by beginning with a random solution, then switching pathways one at a time to find an optimal solution (Glover, 1990).

Previous heuristics that followed a similar model were limited in that they settled for the first optimal solution reached in which another move could not improve the solution (Glover, 1989). In many cases, however, the solutions they discovered were local optima, rather than global optima for the entire domain of solutions (Glover & Laguna, 1997). Tabu, however, "remembers" an optimal solution it has reached, then continues searching for a better solution, even though moving away from the local optimum requires a temporary worsening of solution. To prevent the search from oscillating between two local optima, the heuristic maintains a list of previous moves, and is forbidden from returning to them until a specified number of iterations have elapsed (i.e., these particular moves are "tabu"). The search continues until a prescribed number of iterations have occurred without an improvement in solution (Salhi, 1998). The Tabu algorithm has been shown to be both faster and more effective than previous heuristics in solving complex combinatorial problems (de Amorim, Barthelemy, & Ribeiro, 1992; Malek, et al., 1989; Glover, 1990).

The social network analysis program UCINET 6 (Borgatti, et al., 2002) makes use of the Tabu heuristic in a process called Factions. Factions first divides a directed, polytomous network into a specified number of "clique-like" groups which maximize the number of within-group connections and minimize the number of between-group connections (Borgatti, et al., 2002). Factions calculates a fit statistic based on the hamming distance from an idealized model consisting of all within-group connections and no between-group connections (Borgatti, personal communication, July 30, 2002). It then applies the Tabu algorithm, switching individuals among groups, to minimize the hamming distance from the ideal group. The procedure can be repeated several times from different random starting positions, to verify the robustness of the solution (Borgatti, et al., 1999). In both simulated and real-world networks, Factions and the Tabu algorithm have been demonstrated to be very effective in finding clique-like structures in network data (de Amorim, et al., 1992; Pan, Chu, & Lu, 2000; Moody, 2001).

Although on the surface the factions created by the Tabu algorithm appear very much like clusters, there are some important differences between the two. Like cluster analysis, factions is a disjoint procedure which creates partitions of the entire network. Unlike cluster analysis, however, the factions procedure is a divisive, rather than agglomerative one (Scott, 2000). That is, rather than joining together individual nodes to form larger groups, factions starts with the group as a whole, and attempts to divide it into subgroups. Factions also uses a more advanced algorithm than the cluster analysis procedure, in that the Tabu algorithm can improve on earlier "moves," whereas decisions made by MODECLUS early on are carried through the rest of the procedure.

An additional, and very important, difference should also be noted for clusters and factions compared with cliques and k-plexes. In non-exhaustive cohesive subgroups such as cliques and k-plexes, actors in the network might or might not be found to be members of one or more subgroups, based on their connections with others. In factions and clusters, all actors are placed in exactly one of the network partitions, regardless of their connections with others. This property of disjoint procedures may be a limitation in that the groupings derived are not be as cohesive or interpretable as those found by other methods. However, partitioning the network also provides an important benefit, as it supplies information for every member of the network, rather than only those with especially strong ties to others. Factions and clusters may therefore be seen as complementary techniques with the non-exhaustive methods previously discussed (Everett, n.d.).

Social Position

A final important social networking concept that must be defined is that of social position. Social position is a measure of how similar individuals in a network are in their pattern of ties to other members of the network (Scott, 2000). Concepts like “leader,” “mother,” and “employee” each imply specific types of social connections with others. We might therefore expect to find similar traits or behaviors in individuals with similar patterns of connections with others (Wasserman and Faust, 1994). Social position can be defined in a number of ways. The most common, and restrictive, way of defining social position is the concept of structural equivalence. Automorphic and regular equivalence are less common methods, and will be discussed in less detail below.

Structural Equivalence

Two actors are structurally equivalent if they possess identical connections to and from exactly the same members of the network (Lorrain & White, 1971). In other words, their perceived relationships with others in the network are identical. In terms of social network analysis, structurally equivalent actors are interchangeable within the network (Wasserman & Faust, 1994). In the graph in Figure 1, nodes 12 and 14 are structurally equivalent to one another (each is connected to nodes 11, 13, 15, and 16, and no others).

Although individuals who are structurally equivalent are likely to lie near one another (Wasserman & Faust, 1994), structural equivalence is distinctly different from the concept of cohesive subgroups discussed above. Structural equivalence is interested in a pattern of connections, rather than in the strength of any particular ties. For example, two unpopular or isolated children in a classroom, with no connections to or from any of their classmates, would be considered structurally equivalent, yet not be part of any cohesive subgroup.

In practice, however, it is rare to find perfect structural equivalence (Scott, 2000). Rather, individuals tend to demonstrate varying degrees of equivalence with one another based on the connections to them and from them. A few of the many methods of measuring equivalence will be discussed here.

The measure of structural equivalence most familiar to non-SNA researchers is the Pearson product-moment correlation. A correlation coefficient can be computed between each pair of actors, based on their pattern of connections to others. For two actors *i* and *j* in a sociomatrix, the mean of row *i* (that is, *i*'s connections to others) is denoted $\bar{x}_{i\bullet}$, and the mean of column *i* (others' connections to *i*) is denoted $\bar{x}_{\bullet i}$. The formula (Wasserman & Faust, 1994, p. 368) for the correlation between *i* and *j* is then computed:

$$r_{ij} = \frac{\sum (x_{ki} - \bar{x}_{\bullet i})(x_{kj} - \bar{x}_{\bullet j}) + \sum (x_{ik} - \bar{x}_{i\bullet})(x_{jk} - \bar{x}_{j\bullet})}{\sqrt{\sum (x_{ki} - \bar{x}_{\bullet i})^2 + \sum x_{ik} - \bar{x}_{i\bullet})^2} + \sqrt{\sum (x_{kj} - \bar{x}_{\bullet j})^2 + \sum x_{jk} - \bar{x}_{j\bullet})^2}}$$

Correlations range from –1 (indicating an inverse relation) to +1 (indicating perfect structural equivalence), with 0 indicating no relation.

A related method of determining structural equivalence is by computing Euclidean distance for each pair of actors (Burt, 1987). In a network of size *g*, where x_{ik} is the value of the connection between persons *i* and *k*, the Euclidean distance between persons *i* and *j* is calculated as (Wasserman & Faust, 1994, p. 367):

$$d_{ij} = \sqrt{\sum_{k=1}^{g}[(x_{ik} - x_{jk})^2 + (x_{ki} - x_{kj})^2]}$$

Euclidean distances range from 0 (indicating perfect equivalence) to a maximum of $\sqrt{2(g-2)}$, signifying no equivalence (Wasserman & Faust, 1994).

In many cases correlation and Euclidean distance provide comparable estimates of structural equivalence. However, the two are not identical, and may yield differing results, particularly when patterns of responses are similar but means and variances differ (Cronbach and Gleser, 1953). Although the choice of which to use is not clear-cut (Aldenderfer & Blashfield, 1984), the present study utilizes correlations, which may be preferable when using polytomous self-report ratings of friendship (Wasserman & Faust, 1994; Faust and Romney, 1985).

Automorphic and Regular Equivalence

Two less restrictive measures of social position, besides structural equivalence, should also be noted: Automorphic equivalence, and regular equivalence. Structural equivalence is actually a special case of automorphic equivalence, which is in turn a special case of regular equivalence (Michaelson and Contractor, 1992). Each of these measures will be briefly described here.

Structural equivalence requires that actors be connected in identical ways to exactly the same other actors. In contrast, automorphic equivalence requires only that actors be connected in identical ways to the same *number* of actors, who are themselves automorphically equivalent to one another. Two automorphic actors will therefore have the same indegree, outdegree, and centrality, belong to the same number of cliques, etc. (Wasserman and Faust, 1994). They will be indistinguishable from one another if the labels on the graph are removed (Michaelson and Contractor, 1992).

Like automorphic equivalence, regular equivalence requires that two actors be connected in identical ways to similar actors. However, regular equivalence relaxes these requirements further, and does not consider the number of these connections (Michaelson and Contractor, 1992). Therefore, two regularly equivalent actors will be have similar ties to individuals who are themselves regularly equivalent to one another, but will not necessarily have the same number of these connections (Wasserman and Faust, 1994).

In comparing structural, automorphic, and regular equivalence, consider a hierarchical corporate network. Two mid-level managers would be regularly equivalent if they each supervised one or more workers, and each were supervised by one or more higher-level executives. To be automorphically equivalent, they would need to supervise the same number of workers, and to be supervised by the same number of executives. To be structurally equivalent, they would need to supervise exactly the same workers, and be supervised by exactly the same executives. Note that regular and automorphic equivalence may be computed across networks (two managers in two different corporations may be regularly equivalent, for instance), but that structural equivalence may not (Wasserman and Faust, 1994).

Regular and automorphic equivalence are more recent concepts than structural equivalence, and have been less widely used (Wasserman and Faust, 1994). I will therefore not go into the details of calculating these measures. However, like structural equivalence, both regular and automorphic equivalence may be computed as dimensional measures, and may be partitioned into similar groups using hierarchical cluster analysis and other methods. (For more technical details, see Borgatti and Everett, 1989; Everett, 1985; Wasserman and Faust, 1994).

Personality Applications of Social Network Analysis

Social network analysis has been applied to a variety of social sciences ranging from anthropology to business management. Although it has rarely been utilized specifically in the field of personality perception (Kanfer & Tanaka, 1993), several studies have investigated personality correlates of social network positions.

Kanfer and Tanaka (1993) conducted a round-robin design study investigating the association between various social network variables and ratings on the Big Five personality dimensions. Participants were the 26 members of an undergraduate class, who reported whether or not they had interacted with each of the other participants during the 10 week class. Participants then rated one another on one item from each of the Big Five dimensions. The items rated were: "(a) shy versus outgoing, (b) unkind versus kind, (c) irresponsible versus responsible, (d) insecure versus secure, and (e) unsophisticated versus sophisticated" (Kanfer and Tanaka, 1993, p. 725).

Kanfer and Tanaka (1993) found significant positive associations between adjacency (i.e., whether the target and rater had interacted) and self-other agreement for the traits of outgoing and kindness. That is, targets and raters who had interacted in class tended to agree more on their ratings of how outgoing and how kind the target was.

In addition, Kanfer and Tanaka (1993) found that targets' structural characteristics were associated with overall ratings by all raters. Targets' indegree was positively correlated with ratings of "outgoing" (r=.64) and "secure" (r=.61). Ratings on the item "kind" were positively correlated with both targets' outdegree (r=.37) and centrality (.45).

Finally, Kanfer and Tanaka (1993) found that raters' centrality was positively correlated with increased consensus with other judges. The strength of this association varied with the trait being rated, ranging from a correlation of 0.12 for "shy versus outgoing" to 0.39 for "unkind versus kind." In other words, more central individuals made ratings more in line with the rest of the group than did less central judges, particularly when rating targets' kindness. Taken together, the findings of Kanfer and Tanaka suggest that social network characteristics are reliably associated with perceived personality characteristics.

Centrality

Mehra, Kilduff, and Brass (2001) investigated the interaction between social network properties and the personality trait of self-monitoring. Self monitoring reflects an individual's tendency to monitor and alter his or her behavior based on the social environment. High self-monitors attempt to change their behavior to better fit in with those around them; low self-monitors maintain their patterns of behavior regardless of the social setting (Kilduff, 1992). In a review of the literature, Mehra and colleagues note that high self-monitoring is associated with better performance ratings in the workplace. Social networking studies have also associated higher levels of betweenness centrality with better workplace performance, presumably because more central individuals can control the flow of information across the network (Burt et al., 1998).

Mehra and colleagues (2001) investigated the association between self-monitoring, betweenness, and job performance. They collected data from 102 employees of a small technology firm. Participants identified their friendship relations within the organization, as well as completing a self-monitoring inventory. Results confirmed that higher self-monitoring scores were significantly associated with higher betweenness centrality scores in the friendship network (r^2=.17, p<.05). Further, both self-monitoring and betweenness centrality were independently associated with higher performance ratings by supervisors, over and above their shared variance (Mehra, et al., 2001). These results suggest that certain personality traits may predict one's place in a social network. Moreover, the results indicate that both personality traits and network status predict the way one is perceived by others.

Structural Equivalence

Burt, Jannotta, and Mahoney (1998) investigated the personality correlates of what Burt has termed "structural holes" (Burt, 1992). Structural hole theory is primarily concerned with the control and flow of information through a network. In brief, structural holes are areas of a network in which two or more groups of individuals are only weakly connected to one another. Burt (1992) describes the structural hole as an opportunity for entrepreneurial individuals to control the flow of information between these groups. In many ways the structural hole is similar to Freeman's (1979) measure of betweenness centrality between disconnected individuals. Borgatti (1997) has also noted that many of the measures of structural holes are highly correlated with existing social network measures.

In their study, Burt and colleagues (1998) investigated the ego network structures of 51 students in an MBA program. Each student described the network of his or her current or most recent employment, detailing the strength of relationships with each contact, competitor, and difficult co-worker, as well as the relationships among the contacts. The researchers then computed the network constraint (lack of structural holes) in each participant's network. Participants were also administered an organizational personality inventory, consisting of 252 items related to independence, conformity, submissiveness, and other business-related personality traits.

Burt and colleagues (1998) found a strong association between certain personality traits and the level of structural holes in the participants' networks. Specifically, individuals in the least constrained networks (i.e., the most structural holes) described themselves as independent, seeking to get ahead, and thriving on change. In contrast, individuals with few structural holes in their networks tended to be conforming and obedient, seeking security and stability. In other words, as predicted, individuals who connected two or more disparate groups tended to have entrepreneurial, proactive personalities. Using the 10 most predictive personality items, Burt and colleagues (1998) constructed an index scale and found that the scale predicted more than 50% of the variance in network structure. These findings suggest that there is a strong association between network structure and self-reported personality.

Breiger and Ennis (1979) also investigated the relationships between personality and structural equivalence. The researchers analyzed data from 21 undergraduate students over the course of a 13-week discussion group. At the end of the course, each participant rated each other participant on Bales' (1979) SYMLOG personality measure, which consists of three orthogonal factors (summarized as Dominant/Submissive, Friendly/Hostile, and Task-oriented/Emotionally-expressive). Personality ratings were averaged across raters. Each participant also rated how much he or she liked, disliked, and was similar to each other member of the group. These ratings were used to measure correlational structural equivalence within the network, with actors divided into four clusters using the CONCOR algorithm. Analysis of variance indicated that these four clusters, based on structural equivalence of sociometric ratings, differed significantly on both the Dominant/Submissive and Friendly/Hostile dimensions. In other words, individuals with similar peer-reported personality traits tended to make similar types of ratings of others (Breiger and Ennis, 1979).

Michaelson and Contractor (1992) investigated whether individuals whose network structures are similar are also perceived as being similar by other members of the network. Each week over the course of a 14 week undergraduate discussion class, 18 members of the class reported on the amount of contact they had had with each of the other members during the previous week. This information was collapsed across weeks, to create an overall adjacency network of the class. Structural equivalence measures (both correlations and Euclidean distances) were computed for all members of the class. In addition, automorphic equivalence and regular equivalence values were also derived.

On the final week of class, participants were given a list of all possible dyadic combinations of participants, and asked to rate how similar the members of each dyad were to one another, in terms of perceived social type or social role (Michaelson and Contractor, 1992). Perceived dyadic similarity was averaged across judges, and compared with the measures of structural similarity. Structural equivalence, as measured by correlation and Euclidean distance, was not significantly correlated with perceived similarity (r=.35 and .03, respectively; p=ns). The less restrictive measures of structural position, however, were significantly correlated with perceived similarity. (Automorphic equivalence r=.52, p<.01; Regular equivalence r=.52, p<.01). In other words, actors who had similar patterns of interaction in the network were perceived by their peers to be of a similar social type. This effect seemed to be stronger when only the general pattern of contacts was examined (automorphic and regular equivalence) rather than the actual individuals they were in contact with (structural equivalence).

Judgment Accuracy Applications of Social Network Analysis

In addition to social network correlates of personality, numerous studies have investigated the effects of social network position on accuracy in judgments. However, only a few of these studies have dealt with the judgment of personality per se. Most have assessed whether social network position affected one's ability to judge the social network characteristics of others. I review these studies below, beginning with the fundamental question of whether reports of social network characteristics are even accurate.

Are informant reports of social network position accurate?

In the late 1970s and early 1980s, Bernard, Killworth, and Sailer published a series of influential articles on informant accuracy in social network data (Killworth and Bernard, 1976; Bernard and Killworth, 1978; Killworth and Bernard, 1979; Bernard, Killworth, and Sailer, 1980; Bernard, Killworth, and Sailer, 1982), often referred to collectively as the BKS studies. In these studies, the authors collected interaction data from members of various networks, and compared these reports with objective data collected by outside observers. The results of these studies suggested that informants' reports were inconsistent with the observer data. The authors summed up their findings by saying, "…one consistent and unavoidable conclusion has emerged from our studies of informant accuracy in network data: what people say, despite their presumed good intentions, bears no useful resemblance to their behavior" (Bernard, Killworth, and Sailer, 1982, p. 63).

The conclusions of the BKS studies have been disputed by other social network theorists (e.g., Burt and Bittner, 1981; Romney and Faust, 1982; Romney and Weller, 1984; Kashy and Kenny, 1990). It should also be noted that the BKS studies all dealt with informants' attempts to recall, rank order, or otherwise quantify the specific amount of communication held with other individuals in a given time period, rather than general impressions of how well they knew another individual. Nevertheless, the BKS studies have raised questions about the reliability of network informants, and will be summarized here.

Killworth and Bernard (1976) asked 32 deaf participants (all "elite" members of the Washington, DC deaf community) to rank order how much time he or she spent communicating with each other participant via teletype (TTY). The participants were then asked to retain and document each TTY communication over the next 21 days. The authors acknowledged considerable methodological problems in both the data collection process and the study design itself (Killworth and Bernard, 1976; Bernard and Killworth, 1977). Nevertheless, the results suggested that rankings by informants did not strongly predict actual communications. For instance, in only 29% of the cases was a first-ranked target communicated with most frequently, and in only about half the cases were they communicated with 1st, 2nd, 3rd, or 4th most frequently (Killworth and Bernard, 1976).

Killworth and Bernard (1977) replicated the study conducted by Killworth and Bernard (1976), correcting several criticisms of the earlier methodology. Most importantly, they increased the network size, and asked participants to judge communication amounts retrospectively, in addition to predicting it prospectively as in the previous study. Killworth and Bernard (1977) also collected similar informant and observed data in three other samples: "Ham," a network of ham radio operators; "Office," a small social science research firm; and "Tech," the faculty, graduate students, and secretaries in a university technology education program. For the Ham data, a small network of ham radio operators judged how much on-air time they spent talking with one another, which was compared with the logs they kept. For both Office and Tech data, participants ranked how frequently they interacted with one another during a normal working day. These judgments were compared with those of observers who walked through the workspaces every 15 to 30 minutes, noting conversations among dyads. The authors again compared informants' judgments of scaled or rank-ordered communications with observed scaled or rank-ordered communications. The results again suggested low correlations between actual rankings and observed rankings, with a mean correlation across all studies of .48. The authors conclude that "at best, people can recall or predict (on average) less than half their communication (either amount or frequency)" (Bernard and Killworth, 1977, p. 10). Subsequent reanalysis of the same data sets (Killworth and Bernard, 1979) compared triadic structures based on informant data with those computed from observed data. They again concluded that there was virtually no agreement between the two.

Bernard, Killworth, and Sailer (1980) further reanalyzed three of the networks described by Killworth and Bernard (1977) (the Hams, Office, and Tech networks). They also included a new network ("Frat") consisting of a university fraternity house, through which an observer walked every 15 minutes, 21 hours a day, for five consecutive days (Bernard, et al., 1980). For each network, the researchers computed separate clique structures for the observer data and the informants' data. These cliques were then compared using a variety of computational methods. The authors concluded that the cognitive data was wholly inaccurate and, on average, differed from the observed structure by 160%.

In a final study, Bernard, Killworth, and Sailer (1982) examined the communications of 57 scientists via the EIES system, an early electronic mail network. Participants were asked to recall the number and names of individuals with whom he or she communicated over EIES during specified windows of time, which was compared with logs of actual messages sent. Participants tended to vastly underestimate the number of people they communicated with. As in previous BKS studies, the Bernard and colleagues (1982) concluded that correspondence between recall and reality was poor. The authors did, however, note one encouraging sign in the data. Although individual accuracies were poor, global accuracy was fairly good. That is, the most "popular" (frequently communicated with) individuals based on all informant reports were also the most popular based on observed data. In addition, overall relative positioning of participants in the network (i.e., structural equivalence) correlated significantly with that of observed data. The authors conclude that, at a global, aggregated level, members of the network can provide an accurate "feel" for the network structure, but that individual informants are not accurate (Bernard, et al., 1982).

Responses to the BKS studies

In response to the series of articles by Bernard, Killworth, and Sailer, Romney and colleagues (Romney and Faust, 1982; Romney and Weller, 1984) re-analyzed several of the data sets cited by Bernard and colleagues.

Romney and Faust (1982) re-analyzed the "Tech" data, which Bernard and colleagues (1980) had cited as demonstrating no relationship between informant report and observed interactions. Romney and Faust eliminated several outliers, and rescaled the observational data to better reflect its structure. In addition, they rejected many of the a priori assumptions that had been made by Bernard and colleagues (1980), and instead searched for patterns of regularity and accuracy within the data (Romney and Faust, 1982). Based on the changes in their analyses of the data, the authors concluded that there was, in fact, a strong association between informants' reports and the observed data.

Further, they found that actors who interacted more with one another tended to rate the interactions of others more similarly. Finally, they demonstrated that individuals who interacted more tended to share more accurate knowledge of others' interactions.

In a later study, Romney and Weller (1984) re-analyzed all four of the data sets cited by Bernard and colleagues (1980). In contrast to the conclusions of Bernard and colleagues, Romney and Weller found a strong correspondence between informant report and observational data. The authors also found a strong relationship between an individual's reliability with other raters and the individual's accuracy compared with observational data. More reliable raters were also more similar to one another in their responses than were less reliable raters. Although these conclusions were based on informant report of interactions, Romney and Weller emphasize that their findings can be generalized to any type of informant report in which the "truth" is not known. When multiple informants provide a variety of responses, the informants' reliability predicts their accuracy (Romney and Weller, 1984). These conclusions were further elaborated on by Romney, Weller, and Batchelder (1986), who proposed a formal mathematical model for analyzing the accuracy of informant responses when the "correct" answers are not known. Their cultural consensus theory assumes that there is a fixed "common truth," which members of a culture are each privy to in varying degrees. The amount to which informants agree with one another is a function of how much of the "truth" they share. The cultural consensus model has been widely applied to peer perceptions of both personality traits and social network structures (e.g., Iannucci & Romney, 1994;Webster, Iannucci, & Romney, 2002; Freeman, Romney, and Freeman, 1987)

Kashy and Kenny (1990) have also reanalyzed the BKS data, and found the original authors' conclusions unwarranted. Kashy and Kenny examined four of the BKS data sets (Frat, Office, Tech, and Ham), using the social relations model to partition the data into actor, partner, and relationship effects, which were then compared with the observer reports. The authors found that actor effects were not accurate. That is, the overall level that individuals reported interacting with all members of the group seemed to reflect personal biases, rather than observable reality. Partner effects, the overall amount that others reported interacting with a particular individual, were mixed. Although few significant partner effects were found, this measurement was confounded by the poor validity of the actor effects (Kashy & Kenny, 1990). However, high validity was found for specific relationship effects. That is, if an individual reported interacting with a specific partner, that report was confirmed by both the partner's report and the observers' report. These findings contradict the earlier conclusions of the BKS studies. Further, Kashy and Kenny note that internal validity (informant correspondence with partner) was higher than external validity (informant correspondence with observer report). Barring a systematic bias in informant errors, Kashy and Kenny conclude that the observational data were themselves flawed. The authors conclude, therefore, that informants are quite accurate in reporting their social connections with others.

The Effects of Acquaintance

Many studies have investigated the effect of acquaintance on personality judgment. For example, Funder and Colvin (1988) compared Q-sort personality ratings made by a participant, two of his or her friends, and two strangers who had only seen the participant in a five minute videotape. The results indicated that the consensus among acquaintances was significantly higher than that of the strangers. The authors concluded that acquaintance plays an important role in reliable personality judgments.

It seems reasonable to assume that greater acquaintance with a target will lead to more accurate judgments of personality. However, several studies, particularly those of David Kenny and his colleagues, have called this assumption into question. Kenny, Albright, Malloy, and Kashy (1994) reviewed 32 studies of acquaintance among judges and targets, applying generalizability theory to interpret the studies' findings. The authors found that, in cross-sectional studies, long-term acquaintances showed greater inter-judge consensus than did short-term acquaintances. However, in longitudinal studies, consensus did not increase over time. Based on these findings, the authors concluded that consensus among judges is influenced more by communication among judges than it is by acquaintance with the target. That is, judges with the opportunity to discuss their impressions of targets with one another tend to show greater consensus than do judges without communication among themselves.

Kenny and Kashy (1994) analyzed data collected in the late 1950's, consisting of more than 400 undergraduate students from 19 residential groups. Participants rated themselves and each member of their residential groups, in a round-robin design, on 27 traits. Kenny and Kashy factor analyzed these 27 traits, and combined them into four factors: Obnoxious, Competent, Paranoid, and Naïve. Participants also nominated the five members of their groups who were their closest friends. Using the social relations model, Kenny and Kashy (1994) computed the consensus, assimilation, self-other agreement, and assumed similarity for each target and rater combination. Overall, the authors found substantial effects for consensus, assumed similarity, and assimilation. However, the researchers found relatively little self-other agreement, noting “although people tend to agree with others when rating somebody else, if they are rating themselves they will agree with others to a lesser extent….People see themselves somewhat differently from how both friends and acquaintances see them” (Kenny and Kashy, 1994). The authors suggest that this may be an effect of the mostly-negative traits being rated, as only for the positive factor, Competent, did they find substantial self-peer agreement.

In addition to calculating interpersonal perception traits in general, Kenny and Kashy (1994) also compared the same traits when dyads were mutually-reported friends vs. acquaintances. Across all traits, the mean level of rater-rater consensus was significantly increased by 40% when the raters were friends vs. simply acquaintances. The authors suggest that this might be due to increased communication, overlap, shared meaning systems, or a combination of all three. In addition, targets who were friends were rated significantly more similarly (an increase of 34%) than were targets who were acquaintances. Based on the data, the authors suggest that about half of this effect is based in reality (friends really do tend to be similar) and the other half in fantasy (friends are often observed in the same behavioral contexts, which may lead to erroneous assumptions about their similarity). Finally, although Friendship slightly increased the effects of assumed similarity, there was little consistent effect of friendship on self-other agreement.

The Weighted Average Model

Building on these findings, Kenny (1991; 1994) has constructed the Weighted Average Model (WAM), a theoretical model describing factors that determine consensus and accuracy in interpersonal perception. The WAM proposes that acquaintance is less important in interpersonal perception than other factors which are, themselves, related to acquaintance. Kenny (1994) states that "…there are three different sources of disagreement in the rating of a target by two judges: nonoverlap, different meaning systems, and unique impression" (p. 75). More specifically, Kenny (1994, p. 245) describes 11 components comprising the WAM: (a) acquaintance, (b) overlap, (c) within-judge consistency, (d) similar meaning systems, (e) between-judge consistency, (f) weight of physical-appearance stereotypes, (g) agreement about stereotypes, (h) assumed "kernel of truth" in stereotypes (within a judge), (i) "kernel of truth" in stereotypes, (j) weight of unique impression (extraneous information), and (k) communication.

The WAM is not a statistical model for partitioning variance, but rather a theoretical model which may inform an analysis of the sources of variance in person perception. According to Kenny, the weighted combination of these factors determines the covariance, and therefore consensus, in perception among raters. The Weighted Average Model is quite complex, and a full explanation is beyond the scope of this paper. However, several of the WAM factors are particularly relevant to social network analysis, and are described here in further detail.

"Acquaintance" is the amount of information the perceiver has about the target. Kenny (1994) quantifies this factor by defining it as the number of behavioral acts observed. In the present study the dyadic knowing score can be considered a measure of acquaintance. “Overlap” is the extent to which two judges observe the same set of behaviors by the target. Although two perceivers might have similar amounts of information about a target (i.e., acquaintance), this information may stem from observations of very different situations or behaviors. If two perceivers know an individual from different contexts, their bases for judgments may be very different. Conversely, “similar meaning systems” refers to the extent to which two observers, observing the same behavior, impute the same meaning to the behavior. Even given high overlap, two perceivers may interpret the behaviors they have observed in very different ways. Finally, “communication” is the extent to which perceivers communicate their impressions of the target with each other. Judges who have talked extensively about their opinions of a target may be more similar in their ratings (Kenny, 1994).

The WAM differs from some other models of personality perception in that it incorporates numerous factors, and posits that they are all interrelated in their moderating effects. Rather than assume a linear relationship between acquaintance and consensus or accuracy, Kenny (1994) states that the effects of acquaintance are variable based on the other factors of the model. For instance, when overlap is absent, greater acquaintance does indeed lead to greater consensus. As overlap increases, however, the effects of acquaintance decrease, and become negligible when overlap is high (e.g., Park, DeKay, and Krauss, 1994). Studies which neglect to account for overlap, therefore, do not obtain an accurate picture of the effects of acquaintance (Kenny, 1994). Similarly, communication among perceivers, shared meaning systems, and agreement about stereotypes may all moderate consensus, both alone and in tandem with other WAM factors (Kenny, 1994). In addition, evidence from the social network literature suggests that network structures may influence social cognition more directly. This literature is reviewed in the next section.

Social Cognition Applications of Social Network Analysis

In a review of the literature, Pattison (1994) outlined three broad ways in which the structure of a social network may be related to social cognition. Pattison emphasizes that these three arguments are not mutually exclusive, and may all work together to shape social cognition. However, Pattison's three categories are a useful framework through which to review the relevant literature on network properties and social cognition.

Information Bias

The first broad category Pattison (1994) describes is "Information Bias." Information bias suggests that the relations in social networks act as channels for the transmission of information. One's location in a network determines, in part, the information that one receives, and therefore the information that can be used in making judgments. This argument suggests that individuals in more central positions, or with connections that provide more relevant information, should be more accurate in their judgments.

Carley (1986) ethnographically studied a group of 45 undergraduates as they worked to select a new graduate resident ("tutor") to live on their dormitory hall. Based on the social network data, Carley divided the participants into structurally equivalent groups. She compared these groupings with a series of interviews with the participants regarding their concepts of what constitutes an appropriate tutor. She found that structural equivalence was associated with increased consensus regarding the concept of tutor. However, good consensus also required social ties between two structurally equivalent individuals; simply being equivalent was not enough in itself. In addition, Carley found that tightly cohesive groups (i.e., clique-like structures) tended to have strong consensus within the group, and weak consensus with non-group members. Carley also found that connectivity with the overall network was associated with increased knowledge over time. Individuals with few ties to the group had a good knowledge base at the outset, but this knowledge did not increase, or even decreased, as the process of selecting a tutor progressed. Individuals with many ties to the network, on the other hand, increased the breadth of their concepts over the course of the study, presumably as they were exposed to different ideas from others in the network (Carley, 1986). Carley concluded that the social structure of a network affects the knowledge acquisition of its members. Individuals tight-knit social groups tend to develop similar cognitions, as do those who are structurally similar (Carley, 1986).

Dean and Brass (1985) studied the network structure of 140 members of a company, based on participants' self-reports of those with whom they frequently interacted. Participants also gave ratings of various job characteristics of employment at the company. Overall job characteristics were also assessed by a trained outside observer. Participants in more central network positions gave more accurate ratings of the job characteristics (i.e., corresponded more highly with the outside observer) than did more peripheral employees. The authors concluded that increased social interaction may allow for exposure to an increased diversity of perspectives, giving a basis for more accurate judgments (Dean & Brass, 1985).

Position as Interaction History

Pattison's (1994) second category of relationship between social context and cognition is "Position as Interaction History." This argument suggests that particular social positions are associated with certain types of social interactions. Repeated interactions may lead an individual to develop certain cognitive patterns and biases. The influence of past interactions brings about expectations for future interactions, and may therefore affect social cognition.

Freeman, Romney, and Freeman (1987) observed the members of a nine-week seminar series at an organization, recording both the attendance and the seating patterns of participants each week. Five days after the final meeting, participants were asked to recall which members had attended the final session. Participants were fairly accurate, but both made errors of omission and falsely recalled people who had not actually attended the final meeting. However, participants' errors tended to be in the direction of the overall pattern throughout the seminar series. That is, the individuals they falsely recalled tended to be those who had attended many previous sessions; the individuals they falsely omitted tended to be those who had missed many previous sessions. The authors suggest that, consistent with studies in cognitive psychology, informants' gaps in memory are filled in by knowledge of long-term patterns.

In addition, Freeman and colleagues (1987) note that, though all of the participants were involved in the same organization, some individuals were based in the central facility (the "in-group"), while others were more peripherally involved in the organization, and based out of satellite offices (the "out-group"). The authors compared group status with informant accuracy in recalling attendance at the final seminar meeting. As predicted, in-group members made many more false recalls than did out-group members, whereas out-group members made many more errors of omission. The authors suggest that those who are in-group had greater opportunities to generate internal mental structures about the organization, and therefore were more strongly affected by expectation effects (Freeman, et al., 1987). These findings suggest that members of a cohesive subgroup may tend to generate similar cognitive biases, leading them to make similar informant ratings, particularly about other members of the subgroup.

Kilduff (1990) investigated the relationship of network properties and behavior in 170 MBA students. Participants identified their close friends in the network, as well as individuals they considered similar to themselves. Based on the friendship network, structural equivalence scores (using Euclidean differences) were computed between each dyad. Behavioral similarity was operationalized by comparing the students' patterns in "bidding" for jobs at various organizations recruiting from the school. Kilduff (1990) found that pairs of individuals who either were friends, or who reported being similar to one another, tended to have similar patterns in their organizational bidding, even when variables such as academic concentration and reported job preference were controlled for. In contrast, computed structural equivalence was not related to bidding behavior once other variables were controlled for. Kilduff (1990) suggests that friendship and perceived similarity may be a more valid measure of actual similarity than is structural equivalence.

In a subsequent analysis of the same data, Kilduff (1992) investigated the moderating effects of the self-monitoring trait (see above in the description of Mehra, et al., 2001). Kilduff again compared friendship and bidding behavior, and found a significant correlation between the two (using the non-parametric measure of $\gamma = .16$, $p<.0001$). However, this effect was significantly stronger for high self-monitors ($\gamma = .19$) than it was for low self-monitors ($\gamma = .13$). Taken together, these two studies (Kilduff, 1990; Kilduff, 1992) suggest a relationship between the structure of networks and actual attitudes and behaviors. Kilduff (1992) further suggests that this relationship may be moderated by personality factors.

Krackhardt and Kilduff (1999) measured the network connections of members of four separate organizations (ranging in size from 21 to 33 members). In addition to obtaining individuals' actual connections, the authors asked participants to describe the connections between others in the network. They then measured participants' perceptions of balance among relationships in the network. That is, whether they perceived dyads as being reciprocal in their relationships, and whether they perceived triads as being balanced. The authors found a curvilinear relationship between distance in the network and perception of balance. Participants described relationships close to themselves as being balanced, and described relationships that were distant from themselves as being balanced. Relationships at an intermediate distance, however, were less likely to be perceived as balanced. The authors concluded that participants saw balance in their own relationships because they felt a measure of control over these relationships and could therefore try to balance them. Conversely, participants had little actual information about relationships distant from themselves, and therefore relied on cognitive schemas about transitivity in relationships. For intermediate relationships, however, participants had too much actual knowledge to simply rely on schemas, but were not directly involved enough to change the balance in the relationship (Krackhardt and Kilduff, 1999). These findings suggest that network distance may affect the types of biases and distortions affecting social perception.

Glaskiewicz and Burt (1991) evaluated the social network among the contributions officers of 61 large corporations. Participants identified the officers of the other corporations with whom they had personal contact. Participants also reported their attitudes toward more than 300 nonprofit agencies in the city. The authors compared participants' attitudes toward the nonprofits with the network structure, examining both structural equivalence and relational cohesion. They found that similarities in attitudes were strongly predicted by structural equivalence, but not by relational cohesion. The authors concluded that attitude contagion was shaped by much more by those they perceived to be their peers in the social structure, rather than those with whom they had personal contact.

Structural Balance

The final category proposed by Pattison (1994) is "Structural Balance." Structural balance suggests that one's social cognitions tend to be similar to those with whom one has strong connections. This theory is derived from the social psychological theories of balance (Heider, 1958) and cognitive dissonance (Festinger, 1957). Consider a triad of actors: A, B, and C. A and B are close friends, and A is friends with C, but B strongly dislikes C. This imbalance in the opinions of A and B will place a strain on their relationship. In order to maintain their relationship while reducing the imbalance, A and/or B will change their opinions of C, until their beliefs are compatible (Pattison, 1994). Structural balance, therefore, predicts that two individuals with close network ties will tend to make similar judgments about a third person.

For example, Johnsen (1986) proposed a model of friendship and attitude formation in large groups, based on principles of structural balance. He suggested that as large groups of strangers interact, weak friendships begin to form based on agreement on issues and values. As time passes, and more opinions are discussed, friendships will either increase, if values are similar, or break apart, if values differ strongly. In other words, friendships in large groups initially form on the basis of similar opinions and values. As friendships strengthen, they reach a point where disagreement becomes less likely to damage the friendship, but are more likely to be resolved by altering the opinions such that the friendship leads to agreement. Johnsen's theory, therefore, suggests that friends and cohesive groups will share similar judgments based on both assortative grouping and changes in attitudes to maintain balance.

Casciaro (1998) applied social network analysis to accuracy in social cognition. Participants were 24 employees of an Italian research center, who, given a list of the other employees, identified both their friendships, and the individuals they would go to for advice. Each participant also reported on the friendship and advice networks of each other participant. Participants' social cognition accuracy (i.e., how accurately they predicted others' self-reported friendship and advice dyads) was compared with network structural variables (centrality, employment status, and level in the organizational hierarchy) and personality traits (Need for Achievement, Need for Affiliation, Self-Monitoring, and Extraversion).

Casciaro (1998) conducted regression analyses predicting social cognition accuracy from the structural and personality variables. Relevant to the present study, Casciaro found that raters' centrality scores were significantly associated with their accuracy in predicting others' friendship (R^2=.41) and advice (R^2=.32) networks. That is, individuals who were more centrally located in the network were more accurate in predicting the self-reported social structure of the network. Personality characteristics, particularly need for achievement and need for affiliation were also significant predictors of accuracy. Overall, the combined models incorporating both network structure and personality variables predicted explained more than 55% of the variance in accuracy.

In a related study, Casciaro and colleagues (Casciaro, Carley, and Krackhardt, 1999) conducted further analyses of the previously described data. The researchers again predicted accuracy in perceiving the global friendship and advice network, but also investigated self-other agreement in perceiving one's own friendship & advice network. The researchers investigated the effects of positive affectivity on accuracy and self-other agreement, while controlling for the variables described in the previous study (centrality, hierarchical position, etc.) Positive affectivity was negatively related to self-other agreement of the advice network, but not the friendship network. That is, happier people were less accurate in their perceptions of their own work-related advice networks, but not of their friendships. Positive affectivity was positively related to global accuracy in the friendship network, but not the advice network. In other words, more cheerful people were more accurate at perceiving others' social relations, but not in perceiving others' advice networks.

The Present Study

Studies of perceptions of disordered personality traits have found low self-peer agreement (Klonsky et al., 2002), particularly when peer information is obtained from unselected peer groups (e.g., Oltmanns et al., 1998; Clifton et al., in press). A possible explanation for the latter finding may be differences in the social network positions of both raters and targets. Previous findings suggest that an individual's position in a social network is likely to be associated with both the way that individual perceives other members of the network and the way in which that individual's personality is perceived by others. However, social network analysis has only rarely been applied to personality perception (Kanfer & Tanaka, 1993), and never to the perception of personality pathology. Therefore, in order to understand the relationship between self- and peer-perceptions of personality disorders, a better understanding of the impact of the social network is essential.

The present study investigates the relationship between network structure and self and peer ratings of personality pathology in 21 large groups of peers. Specifically, four questions of interest will be investigated:

1. What personality and demographic factors predict the network structure? (For example, do individuals of the same race or gender tend to form subgroups with one another? Are certain pathological personality traits predictive of an individual's centrality, connectivity degree, or other characteristics of network position?)

2. How do peer ratings within subgroups differ from those between subgroups? (For example, do peers tend to make an increased number of ratings of those in their subgroups?)
3. How does self-peer agreement within subgroups compare to between subgroups? (That is, are the ratings of peers within an individual's subgroup more similar to the individual's rating of him or herself? If so, does this tendency differ by the personality trait being rated? Are certain types of cohesive subgroups better predictors of self-peer agreement than others?)
4. How does rater-rater consensus within subgroups compare with agreement between subgroups? (That is, are the ratings of peers made within a subgroup more similar than those made between subgroups? If so, does this effect vary by personality trait and/or type of subgroup?)

These four questions will serve as a framework for my analyses of how social network structure mediates self and peer perceptions of personality disorders. Previous studies have found evidence that network position is related to normal personality traits. However, no studies to date have investigated the association of network structure with personality pathology. Further, taking social network structures into account in the assessment of personality pathology may improve both self-peer and peer-peer correspondence in ratings. Incorporating social network analysis into assessments could provide a parsimonious way of increasing both reliability and validity in peer ratings, thereby improving our understanding of the relationship between self and peer perceptions of personality pathology.

Methods

Participants

Participants (N=809, 533 male, 276 female) were Air Force recruits who were assessed at the end of 6 weeks of basic training 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. 64% described themselves as white, 16% as black, 4% as Asian, 4% as biracial, 1% as Native American, and 12% 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 personality disorders.

The participants were members of 21 “flights,” groups of 27 to 54 recruits who go through training together. Six of these flights were single-gender male flights, and 15 were mixed-gender flights. Recruits in a given flight tend to know one another quite well. They spend nearly 24 hours a day together, including time training, eating, and sleeping. Recruits’ names are written on their uniforms and are used frequently by their training instructors and in roll calls, such that members of even large flights become very familiar with one another by name. The study was a round robin design, in that each of the 809 participants acted as both a nominator and a potential nominee (“target”) in the peer nomination process. A list of the 21 flights, and a description of the racial and gender composition of each, appears in Table 2.

Table 2
Gender and Racial Distribution of Flights

Flight #	Total *N*	Gender *N* (%)		Race *N* (%)					
		Male	Female	White	Black	Asian	Biracial	Native American	Other
363	36	21 (58.3%)	15 (41.7%)	25 (69.4%)	1 (2.8%)	1 (2.8%)	3 (8.3%)		6 (16.7%)
364	33	20 (60.6%)	13 (39.4%)	22 (66.7%)	6 (18.2%)		2 (6.1%)		3 (9.1%)
365	36	36 (100%)		24 (66.7%)	4 (11.1%)	2 (5.6%)	2 (5.6%)		4 (11.1%)
366	35	35 (100%)		31 (88.6%)	4 (11.4%)				
369	41	22 (53.6%)	19 (46.4%)	26 (63.4%)	11 (26.8%)	2 (4.9%)			2 (4.9%)
370	38	21 (55.3%)	17 (44.7%)	20 (52.6%)	3 (7.9%)	6 (15.8%)	3 (7.9%)	1 (2.6%)	5 (13.2%)
371	35	19 (54.3%)	16 (45.7%)	27 (77.1%)	3 (8.6%)	1 (2.9%)	1 (2.9%)		3 (8.6%)
372	37	19 (51.4%)	18 (48.6%)	18 (48.6%)	10 (27.0%)	1 (2.7%)	3 (8.1%)		5 (13.5%)
395	37	22 (59.5%)	15 (40.5%)	16 (43.2%)	8 (21.6%)		2 (5.4%)		11 (29.7%)
396	34	21 (61.8%)	13 (38.2%)	17 (50.0%)	9 (26.5%)	1 (2.9%)		2 (5.9%)	5 (14.7%)
399	41	41 (100%)		26 (63.4%)	4 (9.8%)	2 (4.9%)	2 (4.9%)		7 (17.1%)
400	41	41 (100%)		29 (70.7%)	5 (12.2%)	2 (4.9%)	3 (7.3%)		2 (4.9%)
401	35	21 (60.0%)	14 (40.0%)	22 (62.9%)	6 (17.1%)	1 (2.9%)	1 (2.9%)	1 (2.9%)	4 (11.4%)
402	35	17 (48.6%)	18 (51.4%)	26 (74.3%)	4 (11.4%)	1 (2.9%)	2 (5.7%)	1 (2.9%)	1 (2.9%)
403	27	11 (40.7%)	16 (59.3%)	19 (70.4%)	1 (3.7%)	1 (3.7%)	2 (7.4%)		4 (14.8%)
404	39	16 (41.0%)	23 (59.0%)	26 (66.7%)	8 (20.5%)	2 (5.1%)			3 (7.7%)
445	36	18 (50.0%)	18 (50.0%)	19 (52.8%)	7 (19.4%)		2 (5.6%)		8 (22.2%)
449	42	42 (100%)		25 (59.5%)	7 (16.7%)	3 (7.1%)	2 (4.8%)		5 (11.9%)
450	45	45 (100%)		29 (64.4%)	10 (22.2%)		1 (2.2%)		5 (11.1%)
452	52	17 (32.7%)	35 (67.3%)	31 (59.6%)	12 (23.1%)	3 (5.8%)		1 (1.9%)	5 (9.6%)
455	54	28 (51.9%)	26 (48.1%)	37 (68.5%)	7 (13.0%)		2 (3.7%)		8 (14.8%)
Total	809	533 (65.9%)	276 (34.1%)	515 (63.7%)	130 (16.1%)	29 (3.6%)	33 (4.1%)	6 (0.7%)	96 (11.9%)

Materials

The Peer Inventory for Personality Disorder

Each participant was administered a computerized battery of measures, which included the self- and peer-report versions of the Peer Inventory for Personality Disorder (PIPD). The self- and peer-report versions of the PIPD each consist of 106 items, 81 of which are lay translations of the 10 DSM-IV personality disorder criteria[1]. These personality disorder 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 Personality Disorders Workgroup. 23 filler items were also included in these measures, based on additional, mostly positive, characteristics, such as “trustworthy and reliable” or “agreeable and cooperative.” The self-report and peer-report versions of items were otherwise identical, with only the relevant questions differing. A copy of the PIPD appears in the Appendix.

Information from large groups can be obtained by either nomination or rating (or, less frequently, rank ordering). In a nomination procedure, peers identify members of the

---

[1]Although there are 79 total DSM-IV Personality Disorder criteria, in the process of translating them into lay language, two criteria were split into separate items for the sake of clarity. The first criterion was the schizotypal criterion “Inappropriate or constricted affect”, which was rewritten 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 was rewritten 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.

group who best demonstrate a particular criterion. This method is more practical for large groups, and demonstrates greater 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 (Kane & Lawler, 1978). In a ratings procedure, peers rate each individual member of the group as to how much he or she demonstrates a trait. Peer ratings are useful in obtaining more general information about each member's level of a trait (Maassen, Goossens, & Bokhorst, 1998), and may have higher test-retest reliability (Asher & Hymel, 1981; Putallaz & Gottman, 1981), although this finding has not been consistently demonstrated (Terry & Coie, 1991). The PIPD 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. Like nominations, this procedure is best able to identify those with an extreme amount of a trait (i.e., those most likely to be personality disordered), but like ratings it provides more detailed information about the perceived level of the trait.

The peer-rating procedure was a round-robin design in which every individual in the group had the opportunity to rate every other member of the group. Items were presented to participants in a quasi-random order. For each item, the participant was shown a list of all members of his or her group, and asked to nominate those who 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. Peer report scales, based on the DSM-IV criteria sets, were calculated by averaging the scores received for the items in each scale, resulting in a dimensional scale ranging from 0 to 3. The scores assigned by each judge on each scale were kept separate for each target, such that in a flight with N members, each person received (N-1) peer-report scores on each diagnostic scale.

Although ratings by individual raters of targets were kept separate for most analyses, in some instances it was useful to combine ratings of a target across all raters, in order to conduct target-level analyses (see below). In these cases, aggregate peer scores were constructed for each target by taking the mean of all raters' ratings for each of the diagnostic scales. Each target therefore had ten aggregate peer scales, ranging from 0 to 3, which corresponded to the ten peer diagnostic scales.

An additional item, "Please rate how well you know each person" was included to investigate how well acquainted participants were with the individuals they nominated. Unlike the other PIPD items, this item was phrased as a rating-type item, such that participants were encouraged to consider their ratings of each group member, rather than nominating only the most relevant member(s). This item was rated using a four-point rating scale ranging from 0 (not at all) to 3 (very well). Responses to the item, called "knowing", were used to construct an affiliation matrix as described below.

Following the peer-report section, all participants completed a self-report version of the same items. Participants were presented with the items in the same order, and asked "What do you think you are really like on this characteristic?" Participants responded using a 4 point scale: 0 (never this way), 1 (sometimes this way), 2 (usually this way), and 3 (always this way). For each personality disorder, the ratings for the relevant criteria were averaged to form a dimensional measure of personality disorder ranging from 0 to 3.

We recently reported (Thomas, Turkheimer, & Oltmanns, 2003) on the psychometric properties and factor structure of the PIPD in two large non-clinical samples (Sample 1 *N*=2111 Air Force recruits, of which the present study uses a subsample; Sample 2 *N*=1536 college students). The inter-rater reliability for peer ratings on the PIPD items (i.e., the median coefficient alpha across groups, calculated across each of the judges for each PD feature) was .74 in the Air Force sample, with values ranging from .90 to .19 (although only three items had values below .5). In the college sample, reliability ranged from .73 to .26, with a median value of .54. The higher reliability in the military sample likely reflects the greater number of participants per group. Factor analysis of the peer report items also demonstrated high correspondence (congruence coefficients ranged from .87 to .97) with factor patterns of widely used self-report models of PDs (Thomas et al., 2003).

Peer-report items were presented to participants in a quasi-random order. For each item, the participant was asked to nominate members of his or her flight who exhibit the characteristic in question. Participants were required to nominate at least one member of the flight for each item before moving on to the next item. A box labeled “This item was difficult” appeared at the bottom of each screen. Participants were instructed that if they felt that it was particularly difficult to nominate someone for a given item, they should nominate someone who best fit the criterion, and check this box. 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.

Data Analysis Software

The data were analyzed using three statistical software packages. UCINET 6 (Borgatti, Everett, and Freeman, 2002) is a widely-used software package for social network analysis techniques. Based on a user-entered adjacency matrix, UCINET compute a variety social network analyses including measures of connectivity, centrality, density, structural equivalence, and cohesive subgroups.

Pajek (Batagelj and Mrvar, 2001) is a software package for the analysis of large networks. It can perform many of the same network analyses as UCINET; however, for this study it was used primarily to generate graphical displays of the network structures.

The SAS statistical program, Version 8.02 (SAS Institute, 2001) was used to conduct the majority of the analyses predicting outcome measures from social network analysis variables. In particular, SAS's PROC MIXED function was used to construct and analyze linear mixed models as described below.

Procedure

Two or three flights at a time were brought to a central testing center at Lackland Air Force Base. Each participant was seated at a separate computer terminal, where he or she gave written informed consent to participate in the study. After giving consent, they first completed a computerized tutorial on how to select items by pointing and clicking using a mouse, before being administered the PIPD measures. The battery took an average of two hours to complete. During this time, participants were instructed not to talk to one another and to raise their hands if they encountered a problem or question. Dividers between workstations prevented participants from seeing the computer screens of those around them.

Data Analysis

Social Network Analysis

For each flight, an adjacency matrix was constructed based on each participant's Knowing score of each other individual. For a flight consisting of N participants, this consisted of an N x N asymmetrical directed matrix of how well (on a scale of 0 to 3) he or she knows each other individual.

This adjacency matrix was analyzed using UCINET to determine characteristics of the social network. The majority of analyses were conducted using the asymmetrical adjacency matrix. Two analyses of coherent subgroups (cliques and k-plexes) required a symmetrical, non-directed matrix. For these analyses, symmetrical and binary matrices were constructed using procedures within UCINET. The matrices were first symmetrized by taking the mean of the pair's knowing scores. That is, if person *i* reported knowing person *j* a 3 (very well), and person *j* reported knowing person *i* a 0 (not at all), the pair were each assigned the mean score of 1.5. This symmetrical matrix, with values ranging from 0 to 3, was then binarized by considering a relationship present if the mean knowing score was greater than 1.5. This cutoff value was chosen to ensure a conservative estimate of relationships, in which at least one actor had to report knowing the other "well" (2) or "very well" (3), and neither could report knowing the other "not at all" (0). The resulting matrix was both symmetrical and dichotomous, with relationships between every pair members of a flight represented as either present or absent.

Using UCINET routines, numerous social network analyses were conducted on each of the asymmetrical adjacency matrices. First, the matrices were analyzed to determine the overall density of each flight, as well as each individual's indegree and outdegree of relationship ties and betweenness centrality within the network. In addition, the correlational structural equivalence for each pair of actors was determined based on their similarity in links to and from other individuals.

The asymmetrical adjacency matrices were analyzed for cohesive subgroups using the Factions routine of UCINET. As described above, factions are constructed by dividing the network into a given number of exhaustive partitions using a Tabu algorithm. For each matrix, two sets of faction groupings were retained, with differing numbers of partitions specified. First, each matrix was divided into exactly two partitions, as the minimum number of meaningful groups that could be created ("2-partition factions"). Second, we retained the greatest number of partitions in which each group contained at least three members. (Three members in a group were necessary to allow for both target and rater variance within the group.) In no case did further increasing the number beyond this maximum number of meaningful partitions increase the number of members in a faction. The maximum number of meaningful factions ("maximal partition factions") retained from each flight is described in Table 3.

Cluster analysis was also used to divide each matrix into exhaustive partitions. The asymmetrical, directed matrix was first transformed to a distance matrix by subtracting each knowing score from 3, such that a knowing score of 3 was represented by a distance of 0, etc. This distance matrix was then submitted to the SAS MODECLUS procedure (Method=1) with a variable-radius kernel. This procedure conducts cluster analysis by starting with each node in a separate cluster, and joining neighbors together to form higher-density clusters (SAS Institute, 2001). The size of the cluster is determined partly by one or more specified smoothing parameters. In this case, the parameter specified, *k*, was the minimum number of individuals allowable in any cluster. As with the Factions procedure, two separate sets of cluster analysis partitions were retained from each flight. First, the partitions resulting from the smallest value of *k* which yielded a total of two clusters were retained ("two partition clusters"). Second, the partitions resulting from the largest value of *k* in which no cluster contained fewer than three members were retained. As with Factions, in no case did increasing the number of clusters beyond this point result in a greater number of members in a cluster. The number of maximum clusters in each flight ("maximal partition clusters") are described in Table 3.

Table 3
Social Network Characteristics of Flights.

Flight #	*N*	Density	Number of Cliques	Number of K-Plexes	Number of maximal factions	Number of maximal clusters
363	36	.971	39	51	3	5
364	33	0.938	31	47	3	4
365	36	0.881	31	37	3	6
366	35	0.876	36	64	4	6
369	41	1.006	53	173	3	6
370	38	1.088	80	180	2	5
371	35	1.147	42	206	2	5
372	37	1.016	41	84	3	8
395	37	1.136	56	294	2	6
396	34	1.183	61	204	3	3
399	41	1.068	89	362	3	4
400	41	1.029	58	166	3	4
401	35	1.175	72	182	2	3
402	35	1.109	49	163	2	6
403	27	1.167	20	26	2	4
404	39	1.281	101	466	2	4
445	36	.981	45	117	2	4
449	42	0.945	61	174	3	6
450	45	1.069	88	414	4	7
452	52	0.996	97	430	3	5
455	54	0.984	122	663	5	6
Mean	38.52	1.050	60.57	214.43	2.8	5.10
Median	37	1.029	56	174	3	5
SD	6.13	0.11	26.89	167.41	0.81	1.3

The symmetrical, non-directed adjacency matrices were also analyzed using UCINET. Based on these matrices, all cliques within a given flight were determined with UCINET, which searches for cliques using Bron and Kerbosch's (1973) algorithm (Borgatti et al., 2002). Cliques were limited to a minimum size of 3 members. A binary measure of whether each dyad shared any cliques in common, and a continuous measure of the number in common were retained for use in analyses.

UCINET was also used to find the k-plexes within each flight. As described above, a k-plex is a subgroup in which each member of the size n subgroup has a direct tie to (n-*k*) members of the subgroup, including him or herself. Although k-plexes are less restrictive than cliques, we used a relatively conservative definition for our k-plexes, requiring a minimum size group of 5 members, with a k-value of 2. In other words, the k-plexes were subgroups of 5 or more individuals, in which each member was directly tied to all but one other member. As with cliques, both binary and continuous measures of the number of shared k-plexes for each dyad were retained for analysis.

Levels of Analysis

Because the participants were divided into flights, and ratings within each flight took place in a round-robin design, analyses can be performed at several levels. In order to minimize confusion, before describing the actual analyses performed, the three levels of analysis will be described briefly.

Flight Level Analyses

At the broadest level, analysis can be performed on entire flights, based on characteristics that distinguish flights from one another, such as the density of ties within the flight. For example, in order to better understand the composition of networks, analysis of variance was used to compare the density of single-sex flight with that of two-sex flights. The network-level variables used in analysis were: (a) network density, (b) number of participants in the flight, (c) number of cliques in the flight, (d) number of k-plexes in the flight, (e) number of maximal partition factions, and (f) number of maximal partition clusters

Target Level Analyses

Narrowing the field of analysis, analyses can be performed on individuals within the flights, based on personal social network characteristics, such as an individual's position in the network. These social network variables can be compared with demographic information, self-reported personality traits, or aggregated peer-reported personality traits using standard techniques such as analysis of variance, regression, and chi-square analyses. The social network variables used in target-level analyses include: (a) connectivity indegree, (b) connectivity outdegree, (c) normalized betweenness centrality, (d) faction membership (2-partitions), (e) faction membership (maximal partitions), (f) cluster membership (2-partitions), and (g) cluster membership (maximal partitions).

Relationship Level Analyses

The greatest complexity in these analyses occurs at the level of relationships among targets and raters. Although there were 809 participants, there were a total of 31314 rater-by-target dyads. In order to accurately analyze multiple flights, each containing raters who rate multiple targets, and targets who are rated by multiple raters, it is necessary to model the data using an application of generalizability theory (Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Shavelson & Webb, 1991). Generalizability theory has been widely used to partition the variance in studies of interpersonal perception and judgment (e.g., Malloy & Kenny, 1986; Kenny, 1994; Shrout & Fliess, 1979; Shrout, 1993).

Generalizability theory proposes that ratings and measurements are imperfect estimates of an individual's universe (true) score (Shavelson & Webb, 1991). A rating on any given occasion will vary from the true score based on any number of factors. Consider an individual who is administered a test to measure verbal IQ. The score that the individual receives can depend on many things. Certainly the individual's "true" verbal ability plays a role, but so do the ability of the specific test to accurately measure verbal ability, the idiosyncratic way the individual interprets the test questions, a noisy testing environment, whether or not the individual has a headache on the day of testing, and any number of other factors.

Generalizability theory offers a way to estimate the individual's true score separately from the other factors. If the individual were administered several different measures of verbal IQ, and each of those measures were also administered to several other people, it is likely that each person would have a different mean score across tests, and that each test would have a different mean score across people. The grand mean is the mean of all of their scores across all of the tests. For any given test, individuals would vary around the mean for that specific test. For any given individual, his or her test scores would vary around the mean for that specific individual. Finally, different individuals might interact differently with specific tests (perhaps the wording of a particular test is more confusing for some people than for others). Generalizability theory, then, would allow the total variance for all persons and tests to be divided into several components: A grand mean, a person effect, a test effect, an interaction between person and test, and residual error. (In the example described, it would not be possible to separate the person-test interaction from the residual error, and the two would simply be lumped together.)

A similar procedure can be used to estimate the variance components for personality perception (Shrout, 1993). If, across multiple flights, multiple raters rate multiple targets, we can estimate the separate effect of flight membership, effect of raters, effect of targets, effect of specific rater-target interactions, and the residual error. More formally, the following model, adapted from Shrout (1993) can be used to conceptualize the rating given by Rater r to Target t:

$X_{tr} = \mu + \alpha_t + \beta_r + \gamma_{tr} + e_{tr}$ where:

X_{tr} = The specific rating given by Rater r to Target t

μ = The grand mean of all raters' ratings of all targets on *d*.

α_f = The expected effect of Flight f

β_t = The expected effect of Target t

γ_r = The expected effect of Rater r

δ_{tr} = The specific impression of Target t by Rater r.

e_{tr} = Random error effects.

In other words, across all raters and targets, there is a grand mean for the rating given by raters. Flights have a tendency to deviate from the grand mean, represented by α_f. Overall ratings of target t have some tendency, in general, to deviate from the mean of the flight, represented by β_t. Rater r has some tendency, in general, to deviate from the mean of the flight, represented by γ_r. In addition to these general effects, Rater r has a specific, idiosyncratic impression of Target t, represented by δ_{tr}. δ_{tr} is generally not distinguishable from random error effects, which are represented by e_{tr}. The combination of all of these effects leads to a specific rating of Target t by Rater r on a given occasion.

The variance of this basic model is defined as:

$$\sigma^2(X_{tr}) = \sigma^2(\alpha) + \sigma^2(\beta) + \sigma^2(\gamma) + \sigma^2(\delta + e)$$

Based on these principles, linear mixed models were constructed for the data, using the MIXED procedure for SAS software. The basic equation specified for these linear mixed models is:

$$y_{tr} = \mu + \alpha_f + \beta_{t(f)} + \gamma_{r(f)}$$

where y_{tr} is the predicted value of a rating, α_f, $\beta_{t(f)}$, and $\gamma_{r(f)}$ are all random effects variables, and $\beta_{t(f)}$ and $\gamma_{r(f)}$ are each nested within α_f.

For each diagnostic category of personality disorder, this basic model was used to estimate the variance components of flight, rater, target, and residual (error) variance, using restricted maximum likelihood estimates. These variances were then used to calculate rater consensus, using the formula:

$$\frac{\sigma^2(\gamma)}{\sigma^2(\gamma) + \sigma^2(\delta + e)}$$

Then, augmented linear mixed models were constructed by adding additional predictors to the compact model, and evaluating the incremental improvement from these predictors (see Singer, 1998; Snijders & Bosker, 1994). For example, adding the target's self report score for the appropriate personality disorder resulted in a model of:

$$y_{tr} = \mu + \theta_t + \alpha_f + \beta_{t(f)} + \gamma_{r(f)}$$

where θ_t is a fixed effect variable of self-reported score for target *t*. By comparing this augmented model with the compact model, an estimate of the effect of self-report scores can be calculated, using the formula:

$$\frac{\sigma^2(\gamma_C) - \sigma^2(\gamma_{SR})}{\sigma^2(\gamma_C)}$$

where $\sigma^2(\gamma_C)$ is the rater variance for the compact model, and $\sigma^2(\gamma_{SR})$ the rater variance for the augmented model with self-report added as a predictor. This provides an estimate, similar to an R^2 statistic, of the decrease in rater variance with the addition of self-report to the model. Taking the square root of this value provides a correlation-type statistic.

Once the effect of self report for all participants is estimated, the same analysis can be performed separately on the basis of categorical social network variables. For example, the compact and augmented models can be estimated when targets and raters are in the same cohesive subgroup, and when they are in different subgroups. Rater consensus and the effects of self-report can be calculated separately for within-group and between-groups, to determine the effects of cohesive subgroups on ratings. The relationship-level social network variables included for study were: (a) structural equivalence (correlational), (b) shared clique membership, (c) shared k-plex membership, (d) shared two partition faction membership, (e) shared maximal partition faction membership, (f) shared two partition cluster membership, (g) shared maximal partition cluster membership.

Results

Personality Scale Descriptive Statistics

Each of the ten self-reported personality disorder scales ranged from a minimum possible value of 0 to a maximum possible value of 3. In practice, because scores consisted of the mean of several items, the range was slightly smaller, with scores ranging from 0 to 2.86 (for paranoid and avoidant scales). Mean scores ranged from 0.24 (antisocial) to 0.71 (OCPD). The medians, means, standard deviations, and range of each of the ten scales is described in Table 4.

The individual dyadic ratings of any given rater for any target also had a possible range from 0 to 3 on each of the ten scales. Mean scores ranged from 0.09 (Avoidant) to 0.13 (narcissistic), as seen in Table 4. For each target, all of his or her peer ratings across the entire flight were also aggregated by taking the mean of all ratings assigned him or her on each scale. As seen in Table 4, these values ranged from 0 to 1.53 (narcissistic), with means ranging from 0.09 (avoidant) to 0.12 (OCPD).

Table 5 depicts the intercorrelations of diagnostic scales within modalities (i.e., self report and aggregated peer report). Intercorrelations within a modality were moderate to large for the majority of scales. For aggregated peer report, correlations ranged from .12 (avoidant and narcissistic) to .88 (schizoid and schizotypal), with an overall mean value of .58 (SD=.21). For self report, correlations ranged from .28 (schizoid and dependent) to .73 (schizotypal and borderline), with an overall mean value of .53 (SD=.11).

Table 4

Descriptive Statistics for Self-Reported and Aggregated and Dyadic Peer-Reported Personality Disorder Scaled Scores.

	Variable	Median	Mean	Std Dev	Minimum	Maximum
Self Report (*N*=809)	Paranoid	0.429	0.485	0.463	0.000	2.857
	Schizotypal	0.300	0.372	0.387	0.000	2.200
	Schizoid	0.571	0.616	0.406	0.000	2.714
	Antisocial	0.143	0.236	0.324	0.000	2.429
	Borderline	0.111	0.263	0.338	0.000	2.333
	Histrionic	0.250	0.369	0.356	0.000	2.250
	Narcissistic	0.200	0.294	0.334	0.000	2.200
	Avoidant	0.286	0.391	0.413	0.000	2.857
	Dependent	0.125	0.252	0.334	0.000	2.500
	OCPD	0.625	0.705	0.399	0.000	2.500
Aggregated Peer Report (*N*=809)	Paranoid	0.066	0.098	0.097	0.000	0.710
	Schizotypal	0.089	0.115	0.091	0.005	0.866
	Schizoid	0.066	0.095	0.100	0.003	1.006
	Antisocial	0.053	0.100	0.132	0.000	1.436
	Borderline	0.060	0.091	0.095	0.000	0.944
	Histrionic	0.075	0.115	0.125	0.000	1.180
	Narcissistic	0.071	0.133	0.169	0.003	1.527
	Avoidant	0.061	0.090	0.097	0.000	0.811
	Dependent	0.058	0.093	0.105	0.000	1.038
	OCPD	0.089	0.119	0.094	0.015	0.754
Dyadic Peer Report (*N*=31314)	Paranoid	0.000	0.097	0.246	0.000	2.714
	Schizotypal	0.000	0.093	0.217	0.000	3.000
	Schizoid	0.000	0.112	0.230	0.000	2.714
	Antisocial	0.000	0.098	0.268	0.000	3.000
	Borderline	0.000	0.089	0.213	0.000	3.000
	Histrionic	0.000	0.112	0.261	0.000	3.000
	Narcissistic	0.000	0.130	0.335	0.000	3.000
	Avoidant	0.000	0.088	0.225	0.000	3.000
	Dependent	0.000	0.091	0.246	0.000	2.625
	OCPD	0.000	0.115	0.226	0.000	2.625

Table 5

Intercorrelations Between PIPD Diagnostic Categories for Self Report and Aggregated Peer Report

	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent	OCPD
					Aggregated Peer Report (N=809)					
Paranoid	--	0.63	0.66	0.81	0.82	0.71	0.77	0.39	0.45	0.61
Schizoid		--	0.88	0.63	0.76	0.47	0.42	0.70	0.63	0.40
Schizotypal			--	0.61	0.83	0.52	0.40	0.78	0.76	0.37
Antisocial				--	0.79	0.76	0.76	0.35	0.49	0.42
Borderline					--	0.74	0.66	0.63	0.69	0.50
Histrionic						--	0.85	0.31	0.48	0.50
Narcissistic							--	0.12	0.25	0.68
Avoidant								--	0.85	0.12
Dependent									--	0.16
OCPD										--
					Self Report (N=809)					
Paranoid	--	0.52	0.69	0.57	0.63	0.53	0.50	0.54	0.49	0.51
Schizoid		--	0.69	0.45	0.51	0.32	0.37	0.36	0.28	0.46
Schizotypal			--	0.59	0.73	0.55	0.52	0.60	0.54	0.52
Antisocial				--	0.72	0.64	0.63	0.42	0.57	0.37
Borderline					--	0.65	0.61	0.58	0.67	0.52
Histrionic						--	0.68	0.45	0.60	0.46
Narcissistic							--	0.42	0.54	0.49
Avoidant								--	0.63	0.47
Dependent									--	0.45
OCPD										--

Note. All correlations are p<.001

Table 6 displays the correlations between self report and aggregated peer report of the diagnostic scales. Correlations both in general and between corresponding scales were low. In general, correlations for corresponding scales were higher than for non-corresponding scales. Corresponding scale Rs ranged from .14 (for both narcissistic and OCPD) to .30 (for avoidant), and all were significant at $p<.001$. Non-corresponding scale Rs ranged from 0 to .26 (for self report schizotypal with both peer report schizoid and peer report avoidant).

Statistical Analyses

As described earlier, four primary questions were investigated in this study. First, what factors predict the network structure of flights? Second, how do peer ratings within subgroups differ from those between subgroups? Third, how does self-peer agreement within subgroups compare with agreement between subgroups? And finally, how does rater-rater consensus within subgroups compare with agreement between subgroups? The results of analyses investigating these questions are described below.

What factors predict the network structure of flights?

Social network analysis of the asymmetrical directed adjacency matrices found strong evidence of variations in adjacency within each flight. The density of the 21 flights, defined as the proportion of directed connections (weighted by rating score) to possible connections, varied from 0.881 to 1.28 (M=1.05, SD=0.11), as seen in Table 3. Each flight was analyzed with the UCINET factions routine, and partitioned into 2 to 7 groups. The number of maximal partition factions, defined as the greatest number of partitions in which each faction contains at least three members, ranged from 2 to 5,

Table 6

Correlations Between PIPD Diagnostic Categories for Self Report and Aggregated Peer Report

	Self Report									
Peer Report	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent	OCPD
Paranoid	**.15**	.13	.14	.13	.13	.06	.13	.00	.00	.07
Schizoid	.13	**.23**	.26	.12	.19	.04	.08	.16	.04	.08
Schizotypal	.15	.17	**.28**	.11	.22	.08	.08	.19	.11	.11
Antisocial	.14	.16	.14	**.19**	.14	.12	.16	-.03	.03	.01
Borderline	.14	.10	.21	.14	**.22**	.11	.11	.12	.09	.06
Histrionic	.05	.04	.06	.09	.07	**.17**	.14	-.08	.01	.02
Narcissistic	.03	.05	.00	.07	.01	.09	**.14**	-.15	-.07	.04
Avoidant	.12	.12	.26	.05	.21	.05	.01	**.30**	.17	.07
Dependent	.12	.09	.22	.07	.19	.11	.05	.23	**.18**	.04
OCPD	.01	.07	.03	.01	-.01	.00	.08	-.08	-.08	**.14**

Note. $p<.001$ for all correlations in which $|r| > .10$

with a median of 3. The distribution of maximal partition factions is displayed in Table 3. The adjacency matrices were also partitioned into clusters using the SAS MODECLUS routine. The number of maximal partition clusters, defined as the greatest number of partitions in which each cluster contains at least three members, ranged from 3 to 8, with a median of 5. The distribution of maximal partition clusters is displayed in Table 3.

Analysis of the symmetrical, binary adjacency matrices revealed numerous cliques within each flight, ranging from 20 to 122, with a median of 56 (SD=26.89). The number of k-plexes in flights ranged from 26 to 663, with a median of 174 (SD=167.41). The distribution of cliques and k-plexes is displayed in Table 3.

The mean rating, across all dyads, for how well a rater knew a given target was 1.041 (SD= 1.017, 95% CI= 1.030-1.052). In order to investigate how acquaintance varied within and between subgroups, the mean reported level of knowing when both rater and target were in the same subgroup were compared with that of dyads in which the rater and target were in different subgroups. Six different analyses were performed, comparing knowing by factions (2-partition and maximal partition), clusters (2- partition and maximal partition), cliques, and k-plexes. The results of these analyses are depicted in Table 7, and indicate that mean knowing levels are considerably higher within subgroups compared to that between subgroups. The effect sizes (Cohen’s d) of differences between within-subgroup and between-subgroup ranged from 0.69 (maximal partition factions) to 2.43 (cliques). Overall knowing level within-subgroup was highest for cliques, followed by k-plexes and maximal partition clusters.

Table 7

Mean Dyadic Knowing Rating by Shared Subgroup (N=31314 dyads)

	Different Subgroup					Same Subgroup					Difference	
	N Dyads	Mean	Std Dev	95% CI		*N* Dyads	Mean	Std Dev	95% CI		Mean	Cohen's *d*
Clique	24334	0.66	0.75	0.65	0.67	6980	2.37	0.66	2.36	2.39	1.71	2.43
K-Plex	18832	0.56	0.72	0.55	0.57	12482	1.77	0.96	1.75	1.79	1.21	1.42
Faction												
Two partition	15277	0.61	0.87	0.60	0.63	16037	1.45	0.98	1.44	1.47	.84	0.91
Maximal partition	16585	0.67	0.92	0.66	0.68	14729	1.46	0.96	1.44	1.47	.69	0.84
Cluster												
Two partition	15109	0.58	0.81	0.57	0.59	16205	1.47	1.00	1.45	1.49	.85	0.97
Maximal partition	24235	0.84	0.94	0.83	0.85	7079	1.73	0.96	1.70	1.75	.88	0.93

At the flight level, analysis of variance was used to compare network density in single-gender flights (*N*=6) with that of two-gender flights (*N*=15). Predicting density from same-gender/two-gender flight status found that two-gender flights were significantly more dense than single-gender flights ($F(1,19)=4.50$, $p<.05$). The mean density for two-gender flights was 1.08 (SD= 0.10, 95% CI: 1.02-1.13). The mean density for single-gender flights was 0.98 (SD=0.09, 95% CI: 0.88-1.07).

A correlational analysis of the 21 flights was performed, examining the relationship among number of participants, density, number of cliques, number of k-plexes, number of maximal partition factions, and number of maximal partition clusters within each flight. There was a very strong correlation between number of cliques and number of k-plexes in a flight (r=0.935, $p<.0001$). The number of participants in a flight was correlated with several social network measures, including number of cliques (r=.81, $p<.0001$), number of k-plexes (R=.80, p<.0001), and number of maximal factions (r=.60, $p<.01$). The full correlation table appears in Table 8.

Table 8
Correlations Across Flights Among Number of Participants, Density, and Number of Subgroups (N=21 Flights)

	N of Flight	Density	Cliques	K-plexes	Maximal Factions	Maximal Clusters
N of Flight	--	-.230	**.808**	**.802**	**.603**	.307
Density		--	.268	.290	**-.561**	-.418
Cliques			--	**.935**	.332	-.035
K-plexes				--	.377	.060
Maximal Factions					--	.396
Maximal Clusters						--

Note: Correlations in bold type are |*p*|<.05

Numerous cohesive subgroups were identified in each of the flights. The number of subgroups, particularly the number of cliques and k-plexes, varied substantially across flights. Flights varied from 20 to 122 cliques (M=60.6, SD=26.9; see Table 3) and 26 to 663 k-plexes (M=214.4, SD=167.4; see Table 3). As described in Table 8, the number of cliques and k-plexes in a given flight was highly correlated (r=.94, p<.05). The high correlation between these two is to be expected, as a clique is essentially a special case of a k-plex. The number of cliques and k-plexes in a flight was also highly dependent on the number of participants in the flight (r=.81 and .80, respectively). Again, this is to be expected, as the number of nodes in a network sets an upper limit on the number of permutations of cliques and k-plexes that can be obtained. Cliques and k-plexes were not significantly correlated (r=.27 and .29, respectively) with the density of the flight. Density reflects the ratio of connections that are present (weighted by the strength of those connections) to possible connections. It is notable that, because density takes into account the number of possible connections in a network, it is uncorrelated with the size of the network (r=-.23, p=ns). It seems, therefore, that the number of clique and k-plex subgroups in a flight is more strongly associated with the sheer number of people in the flight than with the density of ties among them.

The number of maximal partition factions in the flights (that is, the maximum number of factions in which all partitions had at least three members) ranged from two to five (M=2.8, SD=0.8; see Table 3). This number was significantly correlated with the number of participants in the flight (r=.60, p<.05; see Table 8). As with cliques and k-plexes, this association is likely due in part to an upper limit on the possible configurations of participants into subgroups. The number of maximal partition factions was significantly and negatively correlated with the density of the flight (r= -0.56, p<.05). That is, the denser the network of connections within the flight, the fewer the number of viable factions that could be created.

The number of maximal partition subgroups obtained by cluster analysis ranged from three to eight (M=5.1, SD=1.3; see Table 3). Although the selection of maximal partition clusters was identical to that for factions (i.e., the maximum number of subgroups with at least three participants in each), the number of clusters was not significantly correlated with the number of factions (r=.40, p=*ns*; see Table 8). In addition, in all but one flight the number of maximal clusters was greater than the number of maximal flights (the exception, Flight 396, had equal numbers of factions and clusters). Unlike the previous subgroups described, the number of clusters was not significantly correlated with either the number of participants (r=.31, p=*ns*) or the density of the flight (r= -.42, p=*ns*).

For an example of how networks were partitioned into factions and clusters, please refer to Figures 2 through 7. These six figures depict Flight 369, selected as being representative of the overall median in terms of its size, density, and number of sub-groups. The network was drawn using Pajek (Batagelj and Mrvar, 2001), and arranged using a Kamada-Kawai algorithm. Figure 2 portrays the network of valued and directed connections, with arrowheads and small numbers indicating the direction and strength of the ties. Larger numbers next to the nodes indicate the participant number. Figure 3 portrays the same network, with connections made non-directional and dichotomized, retaining only the most robust connections. Figures 4 through 7 depict the valued and directed network (with arrowheads and tie strengths removed) partitioned and color-coded into two partition clusters, maximal partition clusters, two partition factions, and maximal partition factions, respectively.

As can be seen by comparing Figures 4 and 6, partitioning the network into two clusters and factions, respectively, results in similar groupings. However, the groups are not identical, as several nodes in the center of the graph (nodes 29, 8, 4, 24, 10, and 30) are assigned to one partition by cluster analysis and another by the faction procedure.

Figures 5 and 7, which depict maximal partition clusters and factions, respectively, illustrate further differences between cluster analyses and the factions procedure. Figure 5 portrays the maximal partition clusters which, in this flight, identified six clusters. The clusters obtained were quite disparate, with numerous small clusters throughout the network. The six clusters depicted in Figure 5 bear little resemblance to the two in Figure 4, as the clusters were conglomerated by combining individual nodes from the ground up. In contrast, Figure 7 depicts the maximal partition factions which, in this flight, identified three factions. Comparing this figure to the two factions in Figure 6, there is a substantial resemblance between the two. The third faction (represented in Figure 7 by the red nodes) was created by shearing off four nodes from one faction and one from the other. The remainder of the network remains unchanged from the two partition solution in Figure 6. The differences between the maximal partition clusters in

Figure 5 and factions in Figure 7 illustrate the differences in their agglomerative and divisive methodologies, respectively.

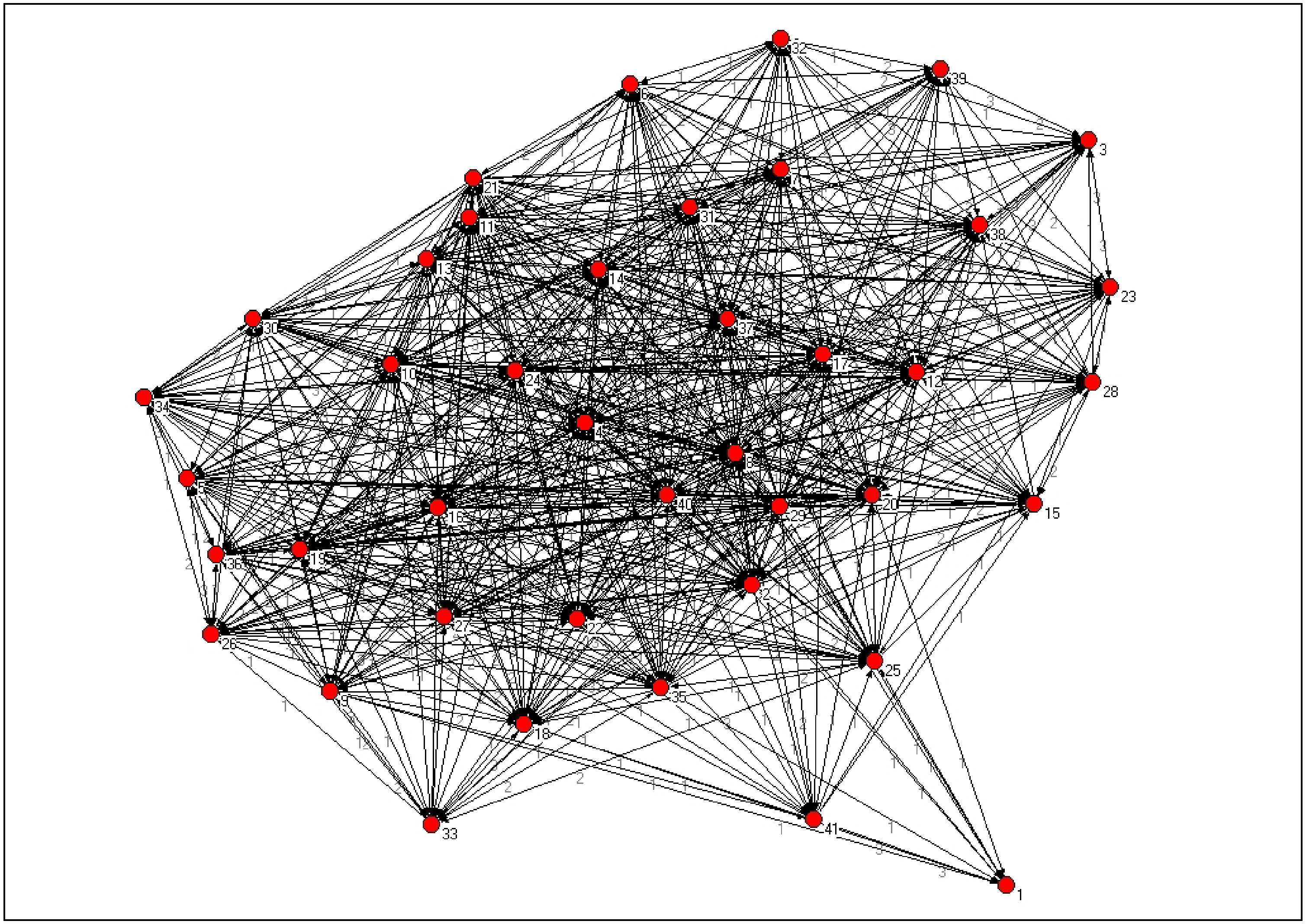

Figure 2. Valued directed graph of Flight 369 with tie values depicted

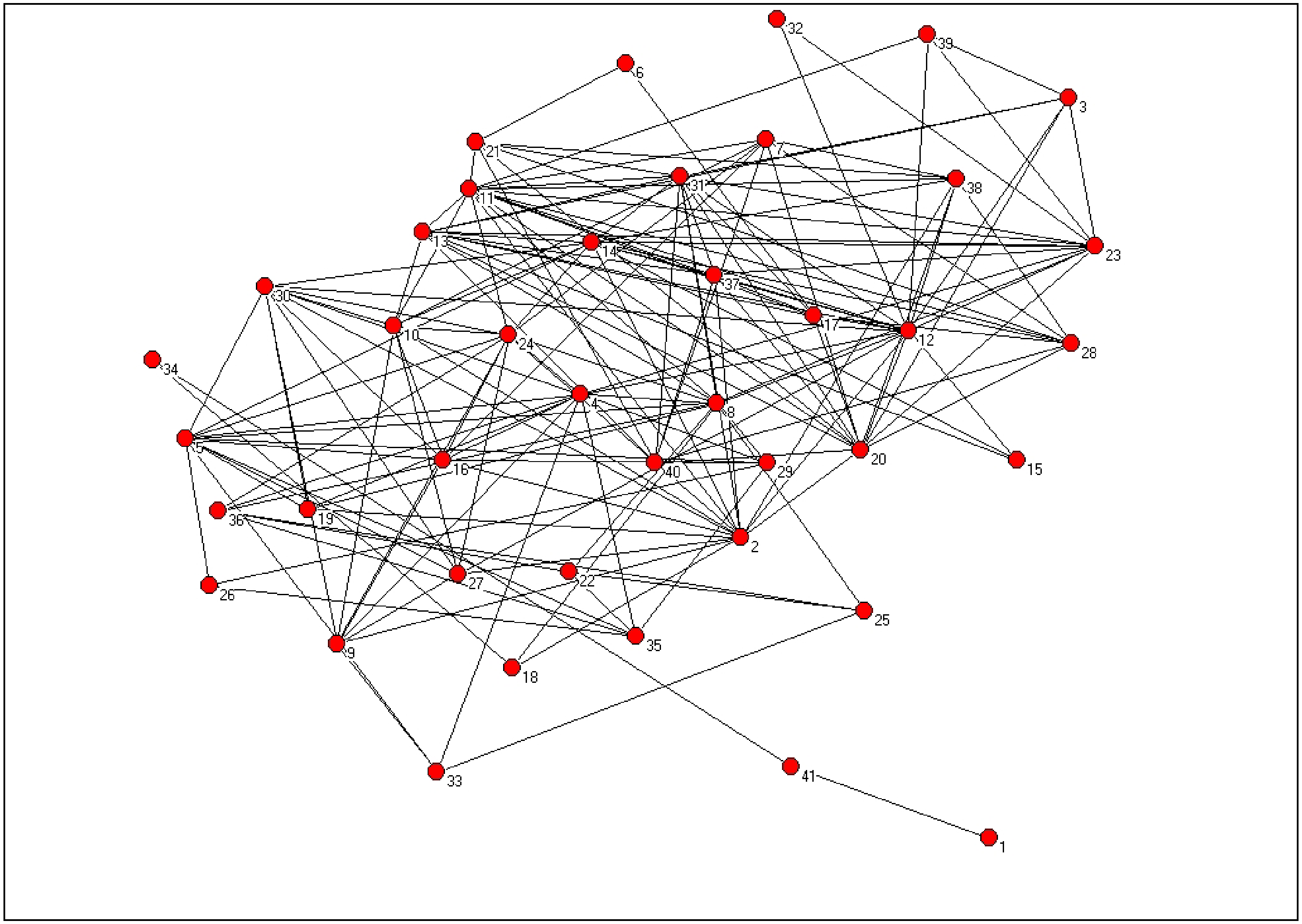

Figure 3. Nondirectional dichotomous graph of Flight 369

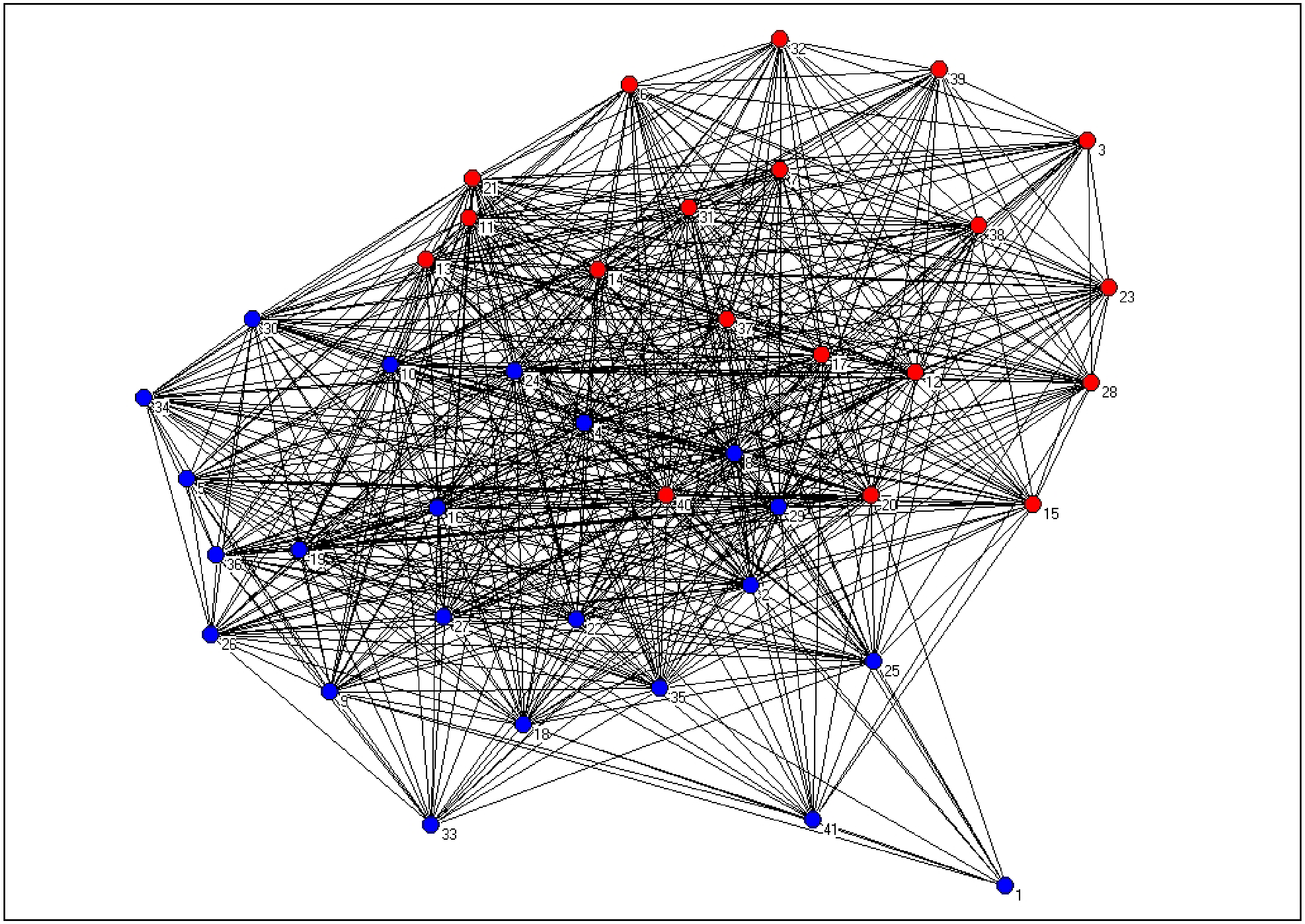

Figure 4. Directed graph of Flight 369 with two partition clusters depicted

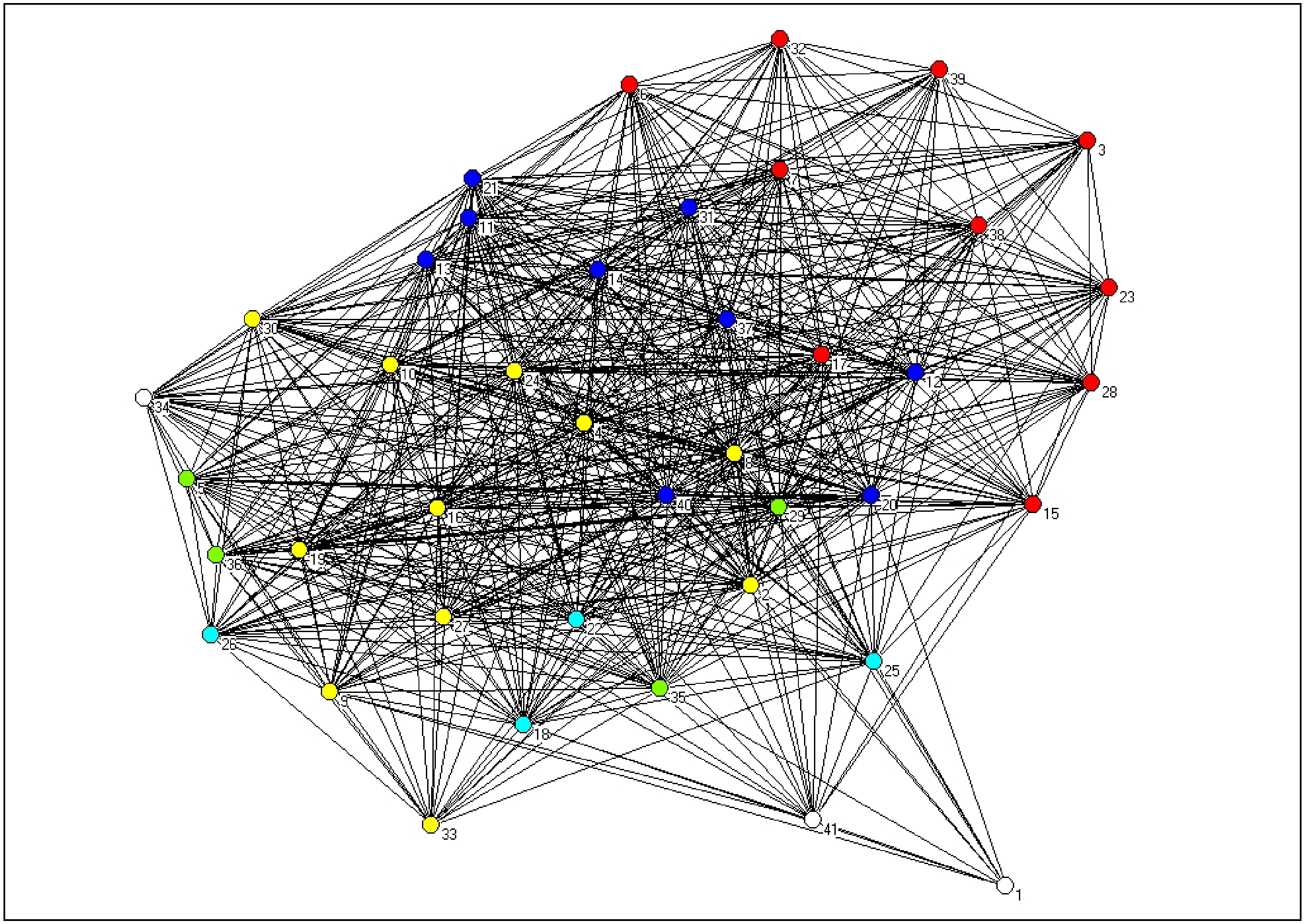

Figure 5. Directed graph of Flight 369 with maximal partition clusters depicted

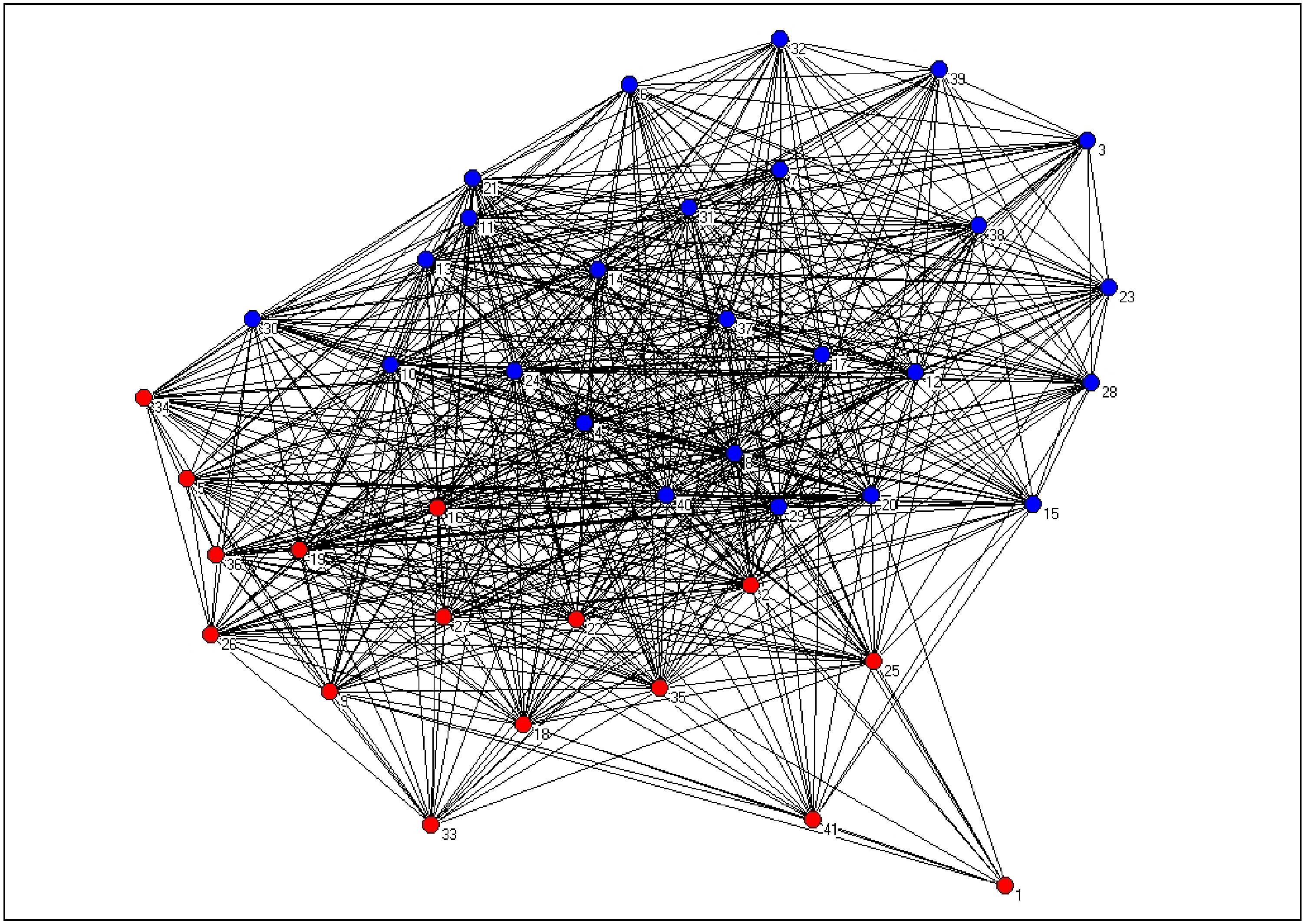

Figure 6. Directed graph of Flight 369 with two partition factions depicted

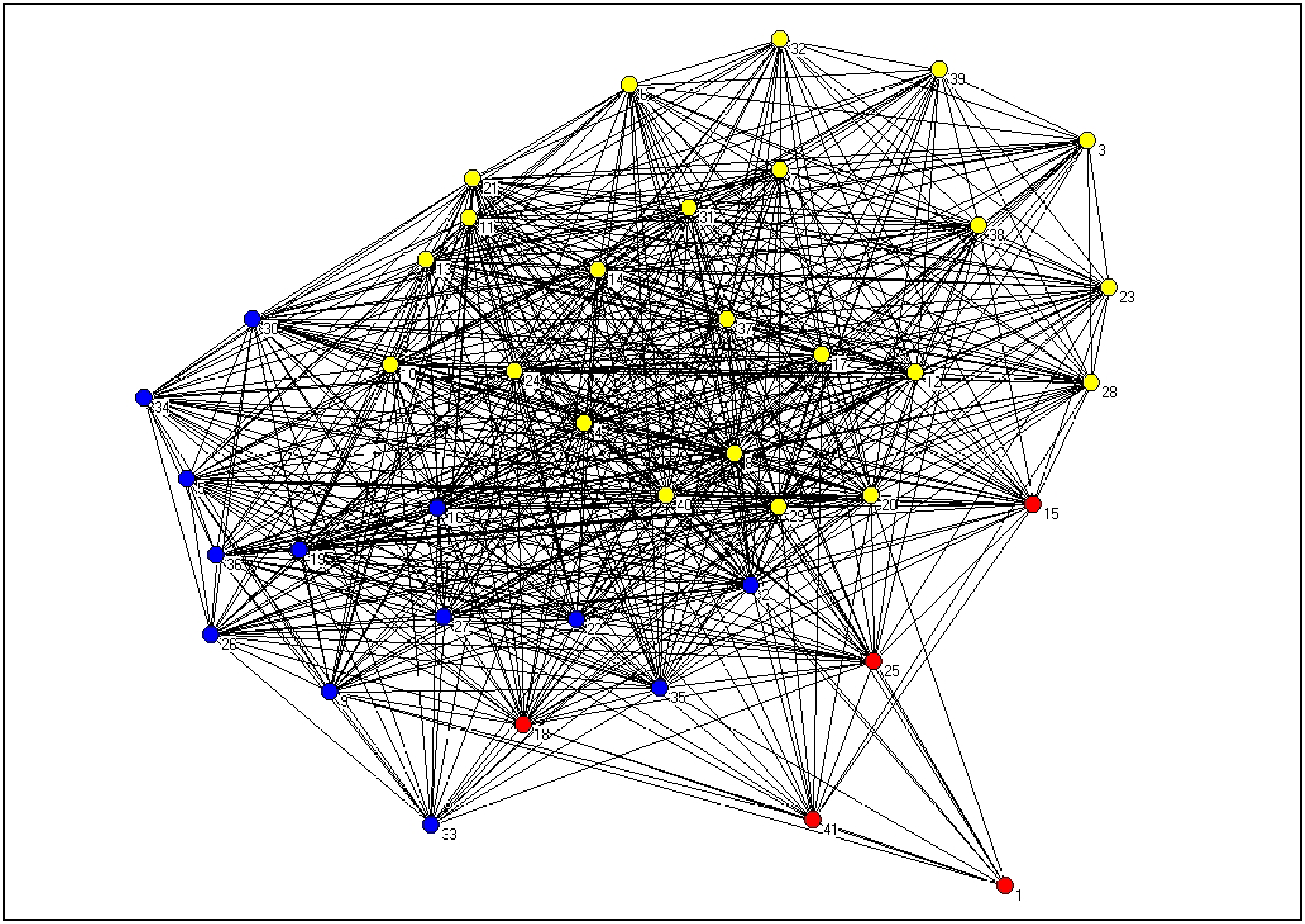

Figure 7. Directed graph of Flight 369 with maximal partition factions depicted

Chi-square analyses were conducted at the dyadic level investigating, for each subgroup type, whether same-group dyads were disproportionately likely to be of the same gender and/or of the same race. The chi-square analyses of gender and subgroups (which included only mixed-gender flights, N=21837 dyads) indicated that there was a very strong association between gender and subgroup. Same-subgroup dyads were disproportionately of the same gender, and different-group dyads were disproportionately of different genders. This held true for all six subgroup types, including cliques (χ^2 [1, N=21837] = 3399.58, p<.0001), k-plexes (χ^2 [1, N=21837] = 4583.36, p<.0001), two partition factions (χ^2 [1, N=21837] = 12986.37, p<.0001), maximal partition factions (χ^2 [1, N=21837] = 8772.93, p<.0001), two partition clusters (χ^2 [1, N=21837] = 14421.07, p<.0001), and maximal partition clusters (χ^2 [1, N=21837] = 5070.41, p<.0001).

Similar chi-square analyses comparing race and subgroup (for the entire sample, N=31314 dyads) revealed significant, but much smaller, tendencies for same subgroup dyads to be of the same race. Significant results were found for cliques (χ^2 [1, N=31314] = 22.95, p<.0001), k-plexes (χ^2 [1, N=31314] = 13.12, p<.001), and maximal partition factions (χ^2 [1, N=31314] = 13.00, p<.001). There was no significant association between race and two partition factions (χ^2 [1, N=31314] = 0.63, p=*ns*), two partition clusters (χ^2 [1, N=31314] = 0.67, p=*ns*), or maximal partition clusters (χ^2 [1, N=31314] = 3.48, p=*ns*).

Also at the dyadic level of analysis, mixed model analyses were conducted predicting structural equivalence and knowing score from whether the pair were of the same race, and (in two-gender flights) whether they were of the same gender. Results suggest that same gender is strongly related to structural equivalence, F(1,21835)=19947,

p<.0001, with an effect size of R^2=0.48. Being of the same gender was also, to a somewhat lesser extent, related to reported knowing score , F(1, 21835)=6224.5, with an effect size of R^2=0.22. Similar analyses using same-race status of dyads as a predictor found significant, but much smaller, effects. Being of the same race predicted structural equivalence, F(1,31312)=13.37, R^2=0.0001, and reported knowing score, F(1,31312)=23.73, R^2=0.001.

In order to investigate how personality characteristics are related to social network variables, across all flights, regression analysis was used to predict an individual's centrality, indegree, outdegree, difference between indegree and outdegree, and symmetric connectivity degree from self-reported personality disorder scales and aggregated peer-reported personality disorder scales. Effect sizes were all relatively small, with r^2 values of prediction ranging from 0 to .07 (predicting outdegree from self-reported avoidant PD). The results of these regressions are described in Table 9.

Table 9

Regression Analyses Predicting Network Position Characteristics From Aggregated Peer-Reported and Self-Reported Personality Disorder Scales (N=809).

		Centrality			Indegree			Outdegree			Indegree-Outdegree Difference		
		R^2	*B*	*p*<	R^2	*B*	*p*<	R^2	*B*	*p*<	R^2	*B*	*p*<
Peer Report	Paranoid	.004	2.60	ns	.001	-5.55	ns	.009	11.64	.01	.007	-17.19	.05
	Schizoid	.004	-2.90	ns	.037	-43.84	.001	.035	-24.01	.001	.008	-19.82	.01
	Schizotypal	.002	-1.83	ns	.019	-28.29	.001	.025	-18.41	.001	.003	-9.88	ns
	Antisocial	.006	2.40	.05	.000	-1.19	ns	.011	9.29	.01	.005	-10.48	.05
	Borderline	.002	2.00	ns	.001	-8.04	ns	.002	4.70	ns	.004	-12.74	ns
	Histrionic	.027	5.41	.001	.004	10.26	ns	.045	19.71	.001	.004	-9.45	ns
	Narcissistic	.013	2.82	.001	.002	4.71	ns	.059	16.77	.001	.011	-12.06	.01
	Avoidant	.012	-4.71	.001	.017	-27.43	.001	.057	-28.86	.001	.000	1.43	ns
	Dependent	.001	-1.32	ns	.005	-13.54	ns	.014	-13.35	.001	.000	-0.19	ns
	OCPD	.015	5.33	.01	.004	-13.35	ns	.024	19.45	.001	.024	-32.80	.001
Self Report	Paranoid	.000	-0.15	ns	.001	-1.51	ns	.000	-0.02	ns	.001	-1.50	ns
	Schizoid	.011	-1.07	.01	.028	-8.52	.001	.008	-2.51	.05	.015	-6.00	.001
	Schizotypal	.004	-0.64	ns	.018	-7.10	.001	.012	-3.26	.01	.006	-3.84	.05
	Antisocial	.006	0.97	.05	.000	0.52	ns	.002	1.61	ns	.000	-1.09	ns
	Borderline	.000	0.16	ns	.001	-1.94	ns	.001	-1.23	ns	.000	-0.71	ns
	Histrionic	.004	0.73	ns	.000	0.12	ns	.005	2.38	.05	.002	-2.26	ns
	Narcissistic	.007	1.01	.05	.000	-0.15	ns	.005	2.53	.05	.002	-2.68	ns
	Avoidant	.024	-1.56	.001	.030	-7.03	.001	.068	-3.94	.001	.005	-3.09	ns
	Dependent	.000	-0.04	ns	.000	-0.41	ns	.004	-2.07	ns	.001	1.66	ns
	OCPD	.001	-0.38	ns	.008	-4.64	.05	.003	-1.62	ns	.004	-3.02	ns

How do peer ratings within subgroups differ from those between subgroups?

This question was investigated through analyses at the rater-target dyadic level. Differences in predicted mean ratings were examined by constructing an augmented linear mixed model, predicting peer rating on a given scale from whether the target and rater shared a mutual cohesive subgroup. The subgroups analyzed in this way were factions (2-partition and maximal partition) and clusters (2-partition and maximal partition), as well as whether the target and rater were in any mutual cliques or k-plexes. In all cases, being in the same subgroup increased the predicted mean rating significantly, although the effect size of the change increase was small, and varied by diagnostic scale and subgroup of analysis. The predicted mean ratings as a function of mutual rater-target cliques and k-plexes, clusters, and factions are described in Tables 11, 12, and 13, respectively.

As seen in Table 10, the effect of mutual subgroup membership on mean ratings varied, both by personality disorder category and by the method of subgroup calculation. Averaged across PD and subgroup type, the mean rating in one's own group was 0.14, while the mean rating outside of one's subgroup was 0.08, with a mean Cohen's *d* of 0.24. The largest effect of subgroup membership (across all PD diagnostic categories) occurred for two-partition clusters (Mean *d* =.32). The smallest effect was seen for cliques (Mean *d* = .18). The effect of subgroup membership on mean peer rating also differed across PD diagnostic category, ranging from a mean *d* of .18 (for avoidant) to .34 (for obsessive-compulsive PD).

Table 10

Mean Dyadic Ratings on Personality Disorder Scales for All Participants and Averaged Across Subgroups (N=31314 dyads).

			Averaged Across Subgroup Types			
	All Dyads		Different Group		Same Group	
Diagnostic Category	Mean	*SD*	Mean	Pooled *SD*	Mean	Pooled *SD*
Paranoid	0.097	0.246	0.069	0.20	0.137	0.29
Schizotypal	0.093	0.217	0.075	0.19	0.117	0.24
Schizoid	0.112	0.230	0.091	0.21	0.141	0.25
Antisocial	0.098	0.268	0.076	0.24	0.129	0.30
Borderline	0.089	0.213	0.065	0.17	0.125	0.25
Histrionic	0.112	0.261	0.083	0.22	0.157	0.31
Narcissistic	0.130	0.335	0.103	0.29	0.169	0.39
Avoidant	0.088	0.225	0.071	0.20	0.112	0.26
Dependent	0.091	0.246	0.065	0.20	0.131	0.30
OCPD	0.115	0.226	0.083	0.19	0.162	0.27

Table 11

Mean Dyadic Ratings on Personality Disorder Scales for Rater-Target Dyads Who Did or Did Not Share a Mutual Clique or K-plex (N=31314 dyads).

	Clique						K-Plex					
	Different		Same		Difference		Different		Same		Difference	
Diagnostic Category	Mean	*SD*	Mean	*SD*	*F*-value	Cohen's *d*	Mean	*SD*	Mean	*SD*	*F*-value	Cohen's *d*
Paranoid	0.084	0.23	0.142	0.29	355.6	0.220	0.074	0.21	0.131	0.29	588.7	0.227
Schizotypal	0.088	0.21	0.108	0.23	231.5	0.090	0.083	0.21	0.108	0.23	382.9	0.115
Schizoid	0.104	0.23	0.138	0.24	107.3	0.148	0.100	0.22	0.130	0.24	346.2	0.131
Antisocial	0.091	0.26	0.123	0.29	83.2	0.116	0.080	0.24	0.125	0.30	295.7	0.164
Borderline	0.076	0.20	0.132	0.26	487.7	0.241	0.068	0.18	0.120	0.25	724.8	0.240
Histrionic	0.096	0.24	0.167	0.32	401.9	0.249	0.083	0.22	0.156	0.31	668.3	0.271
Narcissistic	0.122	0.32	0.158	0.37	31.5	0.103	0.106	0.30	0.167	0.38	237.3	0.179
Avoidant	0.084	0.22	0.105	0.25	122.2	0.091	0.078	0.21	0.103	0.25	297.2	0.107
Dependent	0.078	0.22	0.138	0.31	502.2	0.227	0.070	0.21	0.123	0.29	652.1	0.211
OCPD	0.095	0.20	0.181	0.28	998.0	0.349	0.086	0.19	0.158	0.26	1233.9	0.310

Note. All F-values are F(1,31312). All comparisons are significant at p< .0001.

Table 12

Mean Dyadic Ratings on Personality Disorder Scales for Rater-Target Dyads Who Did or Did Not Share a Mutual Two-Partition or Maximal Partition Cluster (N=31314 dyads).

	Two Cluster						Maximal Clusters					
	Different		Same		Difference		Different		Same		Difference	
Diagnostic Category	Mean	*SD*	Mean	*SD*	*F*-value	Cohen's *d*	Mean	*SD*	Mean	*SD*	*F*-value	Cohen's *d*
Paranoid	0.054	0.17	0.136	0.29	1146.0	0.342	0.082	0.23	0.145	0.30	497.4	0.238
Schizotypal	0.061	0.17	0.123	0.25	950.7	0.291	0.084	0.21	0.127	0.25	391.2	0.185
Schizoid	0.076	0.19	0.145	0.26	912.5	0.304	0.103	0.22	0.154	0.25	314.5	0.202
Antisocial	0.062	0.21	0.132	0.31	775.9	0.264	0.085	0.25	0.135	0.31	312.1	0.181
Borderline	0.050	0.15	0.125	0.25	1335.9	0.360	0.075	0.19	0.136	0.26	622.7	0.265
Histrionic	0.067	0.19	0.154	0.31	1226.5	0.342	0.098	0.24	0.164	0.33	565.6	0.249
Narcissistic	0.086	0.26	0.172	0.39	702.8	0.261	0.116	0.31	0.171	0.40	255.2	0.180
Avoidant	0.059	0.17	0.116	0.26	693.1	0.257	0.078	0.21	0.124	0.27	311.8	0.191
Dependent	0.050	0.17	0.129	0.30	1151.0	0.329	0.076	0.22	0.145	0.32	577.9	0.247
OCPD	0.069	0.17	0.157	0.26	1867.3	0.401	0.101	0.21	0.177	0.27	895.3	0.302

Note. All F-values are F(1,31312). All comparisons are significant at p< .0001.

Table 13

Mean Dyadic Ratings on Personality Disorder Scales for Rater-Target Dyads Who Did or Did Not Share a Mutual Two-Partition or Maximal Partition Faction (N=31314 dyads).

	Two Faction						Maximal Factions					
	Different		Same		Difference		Different		Same		Difference	
Diagnostic Category	Mean	*SD*	Mean	*SD*	*F*-value	Cohen's *d*	Mean	*SD*	Mean	*SD*	*F*-value	Cohen's *d*
Paranoid	0.056	0.17	0.135	0.29	1008.3	0.325	0.066	0.20	0.131	0.29	775.8	0.266
Schizotypal	0.063	0.17	0.121	0.25	801.0	0.268	0.071	0.19	0.117	0.25	672.7	0.213
Schizoid	0.080	0.19	0.142	0.26	795.8	0.273	0.087	0.20	0.139	0.25	631.6	0.225
Antisocial	0.064	0.21	0.130	0.31	638.3	0.249	0.072	0.23	0.127	0.31	520.9	0.204
Borderline	0.055	0.16	0.121	0.25	1081.9	0.315	0.063	0.17	0.118	0.25	912.0	0.254
Histrionic	0.073	0.20	0.150	0.30	971.9	0.300	0.082	0.22	0.146	0.30	712.9	0.246
Narcissistic	0.090	0.27	0.169	0.39	550.1	0.236	0.098	0.28	0.167	0.38	460.2	0.203
Avoidant	0.061	0.18	0.115	0.26	652.7	0.244	0.067	0.19	0.112	0.26	584.7	0.198
Dependent	0.052	0.17	0.129	0.30	1086.1	0.320	0.063	0.19	0.123	0.29	907.8	0.244
OCPD	0.073	0.17	0.154	0.26	1519.5	0.369	0.080	0.18	0.154	0.26	1290.3	0.327

Note: All F-values are F(1,31312). All comparisons are significant at p< .0001.

How does self-peer agreement within subgroups compare to between subgroups?

Self-peer agreement was calculated using mixed linear model analysis at the rater-target relationship level. Compact mixed models were first constructed, predicting each of the 10 peer-reported personality disorder scales. Then, augmented models were analyzed which incorporated self-reported scores as predictors of the corresponding scales. The decrease in rater variance was used to calculate the effect of adding self-report to the model for all rater-target combinations, using the formula described above. This effect, which is similar to an R^2 statistic, describes the amount of correspondence between a given peer's rating and the target's self report for each of the personality disorder scales. Taking the square root of this effect size provides a more interpretable statistic, approximating an estimate of the correlation between self and peer. Across all rater-target dyads and diagnostic categories, the mean self-peer correspondence effect size was equivalent to an R^2 of 0.048, equivalent to a correlation of 0.212. The rater variance in the compact and augmented models, effect size of correspondence, and correlational estimate for each diagnostic category appears in Table 14.

After calculating self-report effects for all participants, rater-target pairings were calculated as a function of whether or not they were in a mutual subgroup. The subgroups used in this way were factions (2-partition and maximal partition) and clusters (2-partition and maximal partition), as well as whether the target and rater were in any mutual cliques or k-plexes. For each cohesive subgroup type, two sets of analyses were performed: one for targets and raters in the same subgroup, and one for those who were not. As above, for each set of analyses, compact and augmented models were constructed, predicting each of the 10 peer-reported personality disorder scales. The effect of self-report was calculated separately for within-subgroup and across-subgroup rater-target pairings. In general, being in the same subgroup improved self-peer correspondence substantially, compared to individuals in different subgroups. The effect size for self-peer correspondence within subgroups, between subgroups, and the difference between the two, are described in Tables 15, 16, and 17. Table 14 summarizes the effect size across all six subgroup methods.

Averaged across subgroup types, the mean improvement in rater-target agreement within subgroups varied greatly based on diagnostic category. It must be emphasized that these are mean effect sizes, averaged across the six types of subgroups used. As will be discussed shortly, there was wider variation in improvement due to subgroup methods than due to PD categories. However, the mean values in Table 14 provide an overview of the variation in PD categories.

As seen in Table 14, self-peer correspondence for narcissistic PD increased from .019 across groups to .021 within groups, an increase of only 11%. Likewise, agreement for schizoid, avoidant, dependent, and schizotypal PDs increased by less than 50%. In contrast, the average effect size of self-peer agreement for histrionic PD more than tripled, from .014 across groups to .042 within groups. Self-peer agreement for OCPD, antisocial, and paranoid PD within subgroups was more than double that between subgroups.

Tables 15 through 17 report the comparisons of self-peer correspondence within subgroups to that between subgroups. Improvement in self-peer agreement differed across subgroup types. In Table 15, self-peer correspondence, averaged across PD categories, increased relatively little within cliques compared to between cliques, averaging only a 24% increase in the amount of variance explained. Self-peer agreement on the histrionic and OCPD scales within cliques was more than double that between cliques. However, several diagnostic categories (paranoid, schizoid, schizotypal, antisocial, and borderline) actually had higher self-peer agreement between cliques than within. A similar pattern of findings occurred when dividing flights into subgroups based on k-plexes. Although self-peer agreement was substantially higher within k-plex than between k-plex for several diagnostic categories (histrionic, and, to a lesser extent, antisocial and OCPD), half of the comparisons found better agreement between k-plexes than within (for schizoid, schizotypal, paranoid, avoidant, and narcissistic).

Partitioning the networks into two groups using cluster analysis improved self-peer agreement more reliably than either cliques or k-plexes, as seen in Table 16. Self-peer agreement in the same cluster was more than double that of different clusters for all diagnostic categories but narcissistic. Paranoid PD, in particular, demonstrated very low self-peer agreement between clusters (R^2=.001), but improved considerably within clusters (R^2=.03).

Compared with two partition clusters, partitioning the flights into the maximum number of viable clusters resulted in less within-cluster improvement in self-peer agreement. Several of the diagnostic categories (narcissistic, schizoid, schizotypal, and avoidant) had worse self-peer agreement within cluster than between. However, despite the lower levels of improvement, note that the within-cluster self-peer agreement for four of the diagnostic categories (paranoid, antisocial, histrionic, and dependent) was higher in maximal partition clusters than in two partition clusters.

Table 17 describes the self-peer agreement when partitioning flights into factions. For two partition factions, in all cases agreement was higher within faction than between. The within-faction improvement was especially robust for OCPD and paranoid PD. Partitioning the flights into the maximum number of factions demonstrated overall lower levels of improvement from between- to within-faction agreement, particularly for dependent PD. However, half of the diagnostic categories (paranoid, schizotypal, antisocial, histrionic, and narcissistic) had higher within-faction agreement in the maximal partition factions than in the two partition factions.

Table 14
Self-Peer Correspondence Effect Size and Square Root of Effect Size for All Participants and Between- and Within-Groups Averaged Across Six Methods of Subgrouping (N=31314 dyads).

			Averaged Across Subgroups			
	All Participants		Between Group		Within Group	
Diagnostic Category	Effect Size	Correlational Equivalent	Effect Size	Correlational Equivalent	Effect Size	Correlational Equivalent
Paranoid	.026	.160	.013	.114	.028	.167
Schizoid	.063	.251	.054	.232	.062	.249
Schizotypal	.089	.298	.053	.230	.079	.281
Antisocial	.039	.199	.021	.145	.046	.215
Borderline	.054	.233	.032	.180	.055	.235
Histrionic	.029	.171	.014	.118	.042	.205
Narcissistic	.023	.152	.019	.138	.021	.145
Avoidant	.106	.326	.082	.286	.103	.321
Dependent	.036	.189	.025	.158	.036	.190
OCPD	.020	.141	.009	.095	.026	.161
Median	.038	.194	.023	.152	.044	.210

Table 15

Effect of Self Report on Peer Report of PD Scales for Rater-Target Dyads That Did or Did Not Share a Mutual Clique or K-Plex, and the Percentage Improvement in Effect Size for Same-Group Dyads.

	Self-Peer Agreement Effect Size					
	Clique			K-plex		
	Different Clique	Same Clique	% Difference	Different K-plex	Same K-plex	% Difference
Paranoid	.028	.015	-46.11%	.032	.021	-34.89%
Schizoid	.093	.056	-40.13%	.101	.059	-41.62%
Schizotypal	.064	.047	-26.47%	.066	.042	-36.06%
Antisocial	.035	.043	23.57%	.022	.046	109.81%
Borderline	.054	.050	-7.93%	.050	.051	1.68%
Histrionic	.019	.051	170.65%	.006	.050	812.68%
Narcissistic	.023	.026	14.25%	.023	.019	-16.63%
Avoidant	.104	.108	4.13%	.110	.088	-20.28%
Dependent	.032	.037	17.21%	.022	.034	51.79%
OCPD	.015	.035	129.07%	.014	.026	82.50%
Median	.034	.045	9.19%	.028	.044	-7.48%

Table 16
Effect of Self Report on Peer Report of PD Scales for Rater-Target Dyads That Did or Did Not Share a Mutual Cluster, and the Percentage Improvement in Effect Size for Same-Group Dyads.

	Self-Peer Agreement Effect Size					
	Two Partition Clusters			Maximal Partition Clusters		
	Different Cluster	Same Cluster	% Difference	Different Cluster	Same Cluster	% Difference
Paranoid	.001	.027	1993.98%	.019	.028	45.63%
Schizoid	.031	.068	120.92%	.054	.042	-22.95%
Schizotypal	.028	.096	238.26%	.082	.066	-18.95%
Antisocial	.016	.040	156.69%	.028	.061	113.76%
Borderline	.008	.059	652.21%	.044	.050	12.36%
Histrionic	.010	.032	234.22%	.020	.050	142.44%
Narcissistic	.016	.020	28.10%	.024	.018	-24.44%
Avoidant	.041	.110	168.91%	.104	.084	-19.75%
Dependent	.011	.034	208.20%	.027	.045	64.54%
OCPD	.003	.023	808.76%	.015	.019	24.53%
Median	.014	.037	221.21%	.028	.048	18.45%

Table 17

Effect of Self Report on Peer Report of PD Scales for Rater-Target Dyads That Did or Did Not Share a Mutual Faction, and the Percentage Improvement in Effect Size for Same-Group Dyads.

	Self-Peer Agreement Effect Size					
	Two Faction			Maximal Partition Factions		
	Different Faction	Same Faction	% Difference	Different Faction	Same Faction	% Difference
Paranoid	-.002	.035	1962.13%	.001	.040	6326.43%
Schizoid	.017	.080	372.85%	.027	.070	156.73%
Schizotypal	.031	.109	254.60%	.047	.111	133.72%
Antisocial	.013	.043	226.39%	.015	.045	204.01%
Borderline	.010	.063	513.26%	.027	.056	103.93%
Histrionic	.015	.029	95.74%	.014	.037	158.66%
Narcissistic	.013	.021	56.78%	.016	.023	39.60%
Avoidant	.061	.114	87.35%	.069	.114	63.46%
Dependent	.022	.038	68.43%	.036	.031	-13.55%
OCPD	.001	.025	4235.92%	.003	.025	754.45%
Median	.014	.041	240.50%	.022	.043	145.23%

How does rater consensus within subgroups compare to between subgroups?

To investigate rater consensus, compact mixed linear models for each of the ten diagnostic scales, across all participants, were first constructed as described above. The rater variance and residual variance (consisting of both idiosyncratic rater-target variance and error variance) were used to calculate the intraclass correlation, as described earlier. ICC(2,1) varied by diagnostic category, ranging from .15 (schizoid) to .25 (narcissistic), with an overall mean of .20. These values, which describe the overall rater consensus for each of the diagnostic categories, are listed in Table 18.

After computing the rater consensus for all rater-target pairings, the same process was done separately for rater-target pairs in the same subgroup, and those in different subgroups. The subgroups used for these analyses were again factions (2-partition and maximal partition), clusters (2-partition and maximal partition), mutual cliques, and mutual k-plexes. Rater consensus, in terms of intraclass correlation, was calculated separately for those raters and targets who shared a subgroup and those who were in different subgroups. In all cases, raters making ratings within their own subgroups had higher levels of agreement than did rater-target dyads which did not share a subgroup. The improvement in ICC for within-subgroup raters compared to across-subgroup raters ranged from 5.3% (maximal partition clusters for schizoid PD scale) to 103.6% (cliques for avoidant PD scale), with a mean improvement of 46.1% across all PD scales and subgrouping methods. Comparing ratings made within-subgroup and between-subgroup, averaged across subgroup types, found higher consensus within groups, ranging from .21 (schizoid) to .31 (narcissistic), with an overall mean of .28. Rater reliability within- and between-subgroups varied somewhat as a function of subgroup type. However, these methodological variations were smaller than those found for rater-target agreement. Mean within-subgroup ICC, averaged across diagnostic category, was highest for maximal partition factions (ICC(2,1)=.29) and lowest for maximal partition clusters (ICC(2,1)=.26). The largest comparative difference for within- and between-subgroup reliability was found by separating participants by cliques. Within-clique ICCs were an average of 59% larger than the corresponding between-clique reliabilities. The rater and residual variances and ICCs within and between subgroups for cliques and k-plexes, clusters (2-partition and maximal partition), and factions (2-partition and maximal partition) are described in Tables 19, 20, and 21, respectively. Table 18 summarizes the rater ICC across all six subgroup methods.

Table 18

Variance and Calculated Intraclass Correlation for Compact Models Across All Rater-Target Dyads, and Mean ICC Averaged Across Subgroup Types (N=31314 dyads).

	All Participants			Mean ICC Across Subgroups		
Diagnostic Category	Rater Variance	Residual Variance	ICC	Different Subgroup	Same Subgroup	% Change
Paranoid	0.008	0.046	.149	.148	.243	64.74%
Schizoid	0.007	0.039	.147	.156	.209	34.07%
Schizotypal	0.009	0.034	.206	.203	.305	50.95%
Antisocial	0.016	0.051	.238	.256	.305	19.87%
Borderline	0.008	0.033	.194	.189	.305	62.30%
Histrionic	0.014	0.046	.234	.245	.308	27.61%
Narcissistic	0.026	0.077	.254	.268	.314	17.19%
Avoidant	0.008	0.038	.177	.161	.287	79.54%
Dependent	0.010	0.045	.175	.165	.298	83.12%
OCPD	0.007	0.034	.171	.200	.242	21.09%
Mean	0.011	0.044	.195	.199	.281	46.05%

Table 19

Comparison of Variance and Intraclass Correlations (ICC) for Peer Report of Personality Disorders Scales for Rater-Target Dyads Within and Between Cliques and K-Plexes.

	Cliques							K-plexes						
	Different			Same			Change in ICC	Different			Same			Change in ICC
Diagnostic Category	Variance Rater	Variance Residual	ICC	Variance Rater	Variance Residual	ICC		Variance Rater	Variance Residual	ICC	Variance Rater	Variance Residual	ICC	
Paranoid	0.006	0.040	.137	0.019	0.057	.246	80.0%	0.006	0.032	.147	0.016	0.058	.218	48.6%
Schizoid	0.007	0.037	.150	0.010	0.037	.211	40.7%	0.006	0.035	.150	0.011	0.040	.222	47.8%
Schizotypal	0.008	0.033	.191	0.015	0.033	.318	66.7%	0.007	0.030	.190	0.017	0.036	.327	71.7%
Antisocial	0.014	0.048	.231	0.024	0.055	.306	32.3%	0.014	0.040	.257	0.023	0.060	.275	6.8%
Borderline	0.006	0.028	.176	0.020	0.041	.326	84.9%	0.005	0.024	.182	0.018	0.040	.304	66.9%
Histrionic	0.011	0.039	.214	0.027	0.062	.307	43.6%	0.008	0.032	.205	0.023	0.061	.279	36.1%
Narcissistic	0.024	0.073	.249	0.035	0.084	.294	18.4%	0.020	0.060	.251	0.038	0.094	.291	16.0%
Avoidant	0.007	0.035	.159	0.019	0.040	.323	103.6%	0.006	0.032	.155	0.018	0.041	.311	100.0%
Dependent	0.007	0.037	.157	0.027	0.059	.316	100.9%	0.006	0.033	.152	0.024	0.053	.309	103.5%
OCPD	0.006	0.028	.182	0.012	0.046	.213	16.6%	0.005	0.024	.178	0.013	0.043	.226	27.4%
Mean			.185			.286	58.8%			.187			.276	52.5%

Table 20

Comparison of Variance and Intraclass Correlations (ICC) for Peer Report of Personality Disorder Scales for Rater-Target Dyads Within and Between Clusters.

	Two Partition Clusters							Maximal Partition Clusters						
	Different			Same			Change in ICC	Different			Same			Change in ICC
	Variance			Variance				Variance			Variance			
Diagnostic Category	Rater	Residual	ICC	Rater	Residual	ICC		Rater	Residual	ICC	Rater	Residual	ICC	
Paranoid	0.004	0.021	.158	0.019	0.056	.252	58.9%	0.007	0.038	.149	0.018	0.059	.238	59.7%
Schizoid	0.005	0.024	.165	0.012	0.045	.206	25.2%	0.007	0.035	.161	0.009	0.044	.169	5.3%
Schizotypal	0.005	0.018	.215	0.017	0.040	.301	40.3%	0.008	0.031	.205	0.014	0.038	.263	28.4%
Antisocial	0.012	0.029	.293	0.028	0.059	.319	8.8%	0.014	0.045	.242	0.024	0.059	.288	19.2%
Borderline	0.004	0.015	.206	0.017	0.040	.297	44.1%	0.007	0.027	.194	0.016	0.043	.276	42.3%
Histrionic	0.008	0.022	.260	0.027	0.056	.324	24.4%	0.011	0.038	.228	0.028	0.059	.324	42.3%
Narcissistic	0.016	0.044	.271	0.046	0.089	.339	24.8%	0.023	0.067	.255	0.042	0.094	.310	21.5%
Avoidant	0.004	0.020	.179	0.017	0.046	.270	50.6%	0.007	0.031	.174	0.016	0.051	.236	36.0%
Dependent	0.005	0.018	.207	0.023	0.056	.288	39.1%	0.007	0.035	.164	0.024	0.062	.282	72.1%
OCPD	0.004	0.015	.202	0.014	0.042	.252	24.7%	0.006	0.028	.186	0.014	0.044	.246	32.2%
Mean			.216			.285	34.1%			.196			.263	35.9%

Table 21

Comparison of Variance and Intraclass Correlations (ICC) for Peer Report of Personality Disorder Scales for Rater-Target Dyads Within and Between Factions.

	Two Partition Factions							Maximal Partition Factions						
	Different			Same			Change in ICC	Different			Same			Change in ICC
	Variance			Variance				Variance			Variance			
Diagnostic Category	Rater	Residual	ICC	Rater	Residual	ICC		Rater	Residual	ICC	Rater	Residual	ICC	
Paranoid	0.004	0.021	.149	0.019	0.056	.255	71.6%	0.005	0.028	.147	0.018	0.054	.249	69.6%
Schizoid	0.005	0.026	.162	0.012	0.044	.216	33.4%	0.005	0.029	.150	0.012	0.042	.228	52.1%
Schizotypal	0.006	0.020	.223	0.017	0.040	.305	37.2%	0.006	0.024	.196	0.017	0.038	.316	61.5%
Antisocial	0.012	0.030	.283	0.027	0.059	.315	11.1%	0.010	0.035	.230	0.028	0.058	.325	41.0%
Borderline	0.004	0.017	.188	0.017	0.039	.310	65.0%	0.005	0.021	.186	0.017	0.037	.318	70.5%
Histrionic	0.010	0.025	.280	0.025	0.055	.311	11.1%	0.012	0.030	.282	0.023	0.053	.305	8.1%
Narcissistic	0.019	0.046	.292	0.044	0.089	.329	12.5%	0.021	0.051	.291	0.042	0.089	.320	10.0%
Avoidant	0.004	0.022	.149	0.017	0.044	.276	85.8%	0.005	0.026	.151	0.018	0.042	.304	101.2%
Dependent	0.004	0.019	.156	0.022	0.057	.281	79.7%	0.005	0.027	.153	0.024	0.052	.311	103.3%
OCPD	0.005	0.017	.232	0.014	0.041	.255	10.2%	0.006	0.020	.223	0.014	0.040	.258	15.5%
Mean			.211			.285	41.8%			.201			.293	53.3%

Discussion

In this study, I applied social network analysis techniques to the study of self and peer perceptions of personality pathology. The goal of the study was to determine whether these techniques, particularly those used to find cohesive subgroups within a network, could help to understand the limitations of peer reports obtained from large groups. Specifically, four questions of interest were examined:

1. What factors predict the network structure of flights?
2. How do peer ratings within subgroups differ from those between subgroups?
3. How does self-peer agreement within subgroups compare with agreement between subgroups?
4. How does rater-rater consensus within subgroups compare with consensus between subgroups?

The overall findings of this study suggest that there are, indeed, identifiable social network structures which can improve our understanding of the relation between self and peer perceptions of personality disorders. Each of the questions of interest will be further explored below.

What factors predict the network structure of flights?

The first question of interest asks whether there are differences in flight characteristics, demographics, or personality traits which contribute to the creation of cohesive subgroups within a network. To begin, flights were examined as a whole to identify factors related to network-level characteristics. Comparing single-gender and two-gender flights found a small but significant difference in network density, in which two-gender flights had a higher mean density than did single-gender flights. One possible explanation for this finding may be that participants in two-gender flights gravitate toward members of their own gender, forming two smaller and tighter-knit subgroups within each flight. This explanation is supported by the results of several analyses of gender at the dyadic level. Being of the same gender was a significant predictor of the dyadic knowing score. Further, chi-squared analyses found that within-group dyads were highly likely to be of the same gender, regardless of the type of subgroup examined. In addition, being of the same gender was a very strong predictor of structural equivalence. That is, two individuals of the same gender tended to have similar connections to and from the same people, as part of a common social circle. Overall, then, gender appears to be a strong predictor of both acquaintance (including direct knowledge and common subgroup membership) and social position (in terms of structural equivalence).

The large association between gender and network characteristics may be related to several factors. It is certainly likely that individuals might feel more comfortable forming friendships with others of their own gender, particularly considering the age of the participants in this study. Other social network studies (e.g., Frank, 1995) have also found tendencies of cohesive subgroups to form among same-gendered participants. An additional factor in the present study may be the structure of the recruits' living situations. Men and women of the same flight are housed in separate barracks by gender, meaning that same-gender participants spend an additional six to seven hours per day together compared with mixed gender dyads. Although much of the time segregated by gender would be spent sleeping, the separate barracks are likely to influence participants' feelings of acquaintance with members of their flights.

Being of the same race was also significantly associated with dyadic knowing, mutual subgroup membership, and structural equivalence. This is consistent with previous research finding that race was associated with cohesive subgroup membership (Frank, 1995) and density of networks (Popielarz, 2000). However, in what may be a positive sign for race relations in the military, the effect sizes of these associations were all considerably smaller than those of gender. Shared race explained less than one percent of the variance in both dyadic knowing and structural equivalence. The largest effect of race occurred for cohesive subgroups, particularly cliques, suggesting that, in this sample, individuals of the same race were more likely to form very tight-knit cohesive subgroups with one another, but were no less likely to be a part of larger subgroups as well.

Characteristics of Cohesive Subgroups

As seen in Table 8, the density of the networks, that is, the number of connections relative to the network's size, was negatively associated with the number of maximal partition factions. In other words, the denser the network of connections within the flight, the fewer the number of viable factions that could be created. This may be a result of the underlying method used to determine factions. As described above, the factions routine utilizes the Tabu algorithm to maximize the value of within-group connections and minimize the value of between-group connections. In a dense network with many interconnections, minimizing the value of between-group connections might be best performed by finding one or two outlying individuals with few connections, and isolating them as individual factions unto themselves. My selection procedure for maximal partition factions required each partition to have at least three members. Therefore, denser networks in which outliers are assigned to individual factions seem to have resulted in fewer maximal partitions than did less dense networks in which more inclusive delineations could be drawn.

One factor which might explain these differences between the correlates of clusters and factions is the agglomerative process of the cluster analysis compared with the divisive process of the factions procedure. Whereas the factions routine tries to separate the full group into a specified number of smaller groups, the cluster analysis routine (using PROC MODECLUS in the SAS program) begins with the individual observations, gradually joining more closely linked nodes to form clusters. In addition, PROC MODECLUS allows the minimum number of members per cluster to be specified, a feature not available in the UCINET factions routine. Taken together, these differences allowed for a greater number of clusters to be created, without the outliers which prevented additional viable factions. This suggests that in a network with many outliers or isolates, cluster analysis may be preferable for identifying multiple usable partitions.

The cohesiveness of the various methods of subgroups can be evaluated by comparing the dyadic knowing score within and between each. As seen in Table 7, the mean knowing score (the rating of how well a rater knows a target, ranging from 0 to 3) was considerably higher for rater-target dyads within a subgroup than for those who did not share a mutual subgroup. This is certainly to be expected, as the subgroups were originally created based on the adjacency matrix of knowing scores. However, as Table 7 indicates, the various types of cohesive subgroups differed in their ability to identify subgroups with higher mean levels of knowing. By far the greatest difference between within-group knowing and between-group knowing occurred for cliques. That is, individuals who were in cliques together reported knowing one another considerably better than those in other types of cohesive subgroups. The fact that cliques contained the fewest within-group dyads (6980) likely plays a role in maximizing the mean knowing score within group. (However, note that maximal partition clusters had only slightly more within-group dyads, 7079, with substantially less differentiation of between- and within-group knowing scores.) More importantly, cliques were derived from a non-directed binary matrix, in which connections were considered present only if the mean knowing score for the dyad was greater than 1.5. This fact, coupled with the restrictive nature of the clique procedure, seems to create subgroups with considerably higher mean levels of knowing within the groups than did any other subgrouping procedure. However, as will be discussed shortly, these increased mean knowing levels did not translate into increased benefits to personality assessment.

K-plexes were the next best able to differentiate subgroups with high levels of within-group acquaintance, whereas maximal partition facets differentiated mean knowing scores the least well. Maximal partition clusters identified a relatively small number of same-group dyads, but, as previously noted, did not differentiate mean knowing scores as effectively as did cliques. Overall, however, all six methods of determining cohesive subgroups identified groups within the network who reported knowing one another better than did individuals who were not in the groups, as in no case was there an overlap in the 95% confidence intervals.

Personality Traits

I next examined social network characteristics at the individual level, to see whether these characteristics are predicted by self- or peer-reported personality traits. Regression analyses were performed, predicting individuals' network centrality, indegree, outdegree, and the difference between indegree and outdegree, from self-reported and aggregated peer-reported PD scales. The results of these analyses are reported in Table 9. Although many of the personality scales turned out to be significant predictors of network characteristics, the effect sizes of these regressions were fairly small. The largest R^2 value reported in Table 9 is only .07 (equivalent to a correlation of .26), predicting outdegree from self-reported avoidant PD. Although this association between avoidant PD and outdegree is small, it is approximately the same size as the association between self-reported avoidant PD and peer-reported avoidant PD. In addition, despite their small effect sizes, the significant predictors are generally in the expected direction, and their interpretation provides some objective evidence of subjective personality traits.

Centrality

Centrality was significantly and negatively associated with peer-reported avoidant, and self-reported avoidant and schizoid PD scales. Betweenness centrality is a measure of an individual's importance in the network, and his or her "interpersonal influence" with others (Wasserman and Faust, 1994, p. 189). In addition, low betweenness (i.e., high structural constraint) has been associated with conformity, obedience, and a need for security and stability (Burt, et al., 1998). It is no surprise, therefore, that avoidant and schizoid personality disorder traits would be negatively associated with centrality. These personality disorders are defined by fear of (avoidant) or disinterest in (schizoid) interpersonal relations. The regression analyses indicate a small but significant tendency for these individuals to remain on the periphery of networks, rather than being in the thick of things.

In contrast, centrality was significantly and positively predicted by peer-reported histrionic, narcissistic, antisocial, and OCPD, and self-reported narcissistic and antisocial PD scales. It is particularly fitting that histrionic PD is associated with centrality, as one of the essential features of histrionic PD is feeling "uncomfortable in situations in which he or she is not the center of attention" (APA, 1994, p. 657). It is interesting that peer-reported, but not self-reported, histrionic PD was a significant predictor of centrality, suggesting that perhaps peers are more accurate judges of this particular tendency.

The positive association of narcissistic and antisocial PDs with centrality is also notable, and consistent with prior research on normal personality. Highly central actors connect otherwise unconnected actors, moving between social groups and acting as "gatekeepers" for social exchange (Freeman, 1979). Burt and colleagues (1998) have found strong associations between high betweenness (actually, low structural constraint) and "entrepreneurial" personality traits such as independence, thriving on change, and using one's advantageous position to get ahead. Normal levels of these traits could be highly adaptive in a social network. However, maladaptive expressions of these traits could be manifested as antisocial and narcissistic personality disorders, as both are marked by manipulativeness and exploitation of others (APA, 1994).

The small positive association between peer-reported OCPD and centrality is less clear, but warrants brief discussion. Based on anecdotal evidence noted in interviews with participants, there is some question as to raters' interpretation of the peer-report items for the OCPD scale. In the Air Force, where hard work and attention to detail are emphasized, peer-reported OCPD may instead tend to identify the individuals given supervisory roles in the flight, rather than individuals who fit the DSM-IV definition of OCPD. Certain individuals in each flight were selected by their training instructors for additional responsibilities. Because their responsibilities included delegation of tasks to other members of the flight, these individuals often seem to have been singled out by peers as being overly devoted to work. The fact that peer-reported, but not self-reported, OCPD predicts centrality suggests that this may be the case in the present study as well.

Indegree

Indegree is a measure of number and strength of connections *to* an individual *from* others. In the present study, indegree was significantly and negatively predicted by both peer- and self-reported schizoid, schizotypal, and avoidant PD scales. That is, higher scores on these PD scales were associated with being less well known by others. The implications of these results are clear, as the DSM-IV definitions of these PDs all include an absence of close relationships. This finding corroborates the findings of Kanfer and Tanaka (1993), who also noted that targets with decreased indegrees were described by peers as less outgoing and less secure, traits often associated with avoidant PD. In addition, it is particularly interesting that the peer reported PDs were predictors of decreased indegree, as it suggests that, even though fewer raters reported knowing these individuals well, raters still singled out these individuals as targets for PD nominations. As a whole, the fact that these PDs are related to indegree in a way consistent with prior research provides evidence of network position consequences related to both self- and peer-reported personality traits.

Outdegree

Outdegree is the counterpart to indegree, and quantifies the connections *from* an individual *to* others. It is essentially one's self-reported degree of connection, whereas indegree is one's peer-reported connection. Contrary to Kanfer & Tanaka (1993), who found few correlates of outdegree, the present study revealed several significant associations with both self- and peer-reported PD traits. As with indegree, self- and peer-reported schizoid, schizotypal, and avoidant PD scales were all negatively associated with outdegree. As would be expected from their DSM-IV definitions, individuals with higher levels of these traits report fewer close acquaintanceships. In addition, peer-reported histrionic, narcissistic, and antisocial traits were all positively associated with increased outdegree. These PDs are all associated, either primarily or peripherally, with being outgoing and gregarious (APA, 1994), which could certainly lead to increased reported acquaintanceship with others. It is particularly notable that outdegree, a self-reported measure, was associated with peer-reported personality scores, providing evidence of the validity of peer-reported PD scales.

The positive association of outdegree with peer-reported paranoid PD, however, is difficult to interpret. Paranoid PD is characterized by mistrust and suspicion of others, including an unwillingness to form trusting relations (APA, 1994). It is curious, then, that these individuals would report greater acquaintanceship with others. One possible explanation for this finding comes from an earlier study of the PIPD (Clifton et al., in press). This study found that individuals who were described by peers as paranoid described themselves as hot-headed and angry. In the present study, the positive association with outdegree provides further evidence that perhaps those described as paranoid by peers do not, in fact, feel particularly mistrustful of others.

Indegree-Outdegree

Finally, the difference between an individual's indegree and outdegree was predicted from self and peer reports. Indegree-outdegree can be thought of as a measure of an individual's expansiveness bias, the tendency to over-report or under-report one's interactions with others, compared to others' perceptions of those interactions (Field & Carter, 2002). A negative B weight for the predictor indicates that greater values of the personality trait scale were associated with increased over-reporting of one's social ties. This was the case for all of the significant predictors, consisting of self-reported schizoid and schizotypal scales, and peer-reported OCPD, narcissistic, schizoid, paranoid, and antisocial scales. The factors discussed for both indegree and outdegree separately apply here as well. However, it may be that the traits measured by these scales are distancing to others, although the individual is unaware of the effect his or her behavior has on others. A classic example of this is narcissism, in which the individual overestimates his or her own worth, while at the same time alienating others with his or her behavior. Interestingly, although peer-reported narcissism was significantly associated with indegree-outdegree, the largest effect was for peer-reported OCPD. This may be another example of the possible confounding of OCPD traits with leadership within the flight. If so, it may reflect individuals with supervisory roles being unaware

How do peer ratings within subgroups differ from those between subgroups?

We have thus far confirmed that there are, indeed, subgroups of participants within each flight, and that individuals' personality traits, gender, and race are, at least in part, related to their positions within their networks. The remaining analyses deal with the differences in ratings made by peers of those within their subgroups and those in different subgroups. For all of the following analyses, six different methods were used to determine cohesive subgroups within each flight. The six methods were: cliques, k-plexes, two partition clusters, maximal partition clusters, two partition factions, and maximal partition factions, all of which will be collectively referred to as "subgroups."

The next question of interest concerns whether peers are more likely to nominate those within their own subgroups. That is, does the mean level of peer rating differ within and between groups? The mean values of ratings within subgroups were compared with those between subgroups. The results of these comparisons, in Tables 10, 11, 12, and 13, indicate that there is a highly significant tendency for the ratings within a subgroup to be higher than those between subgroups, with effect sizes ranging from small to moderate. The largest mean effect size was found for Two Partition Clusters; ratings made within-cluster were on average more than twice as large as those between-cluster, with a mean effect size (across diagnostic categories) of 0.32

Due to the nomination procedure used to obtain peer ratings, the increased mean ratings within subgroups represents both a greater number of nominations received, and a higher average rating on those nominations. Participants were instructed to nominate at least one member of the flight for each item. Although they were free to nominate any member of the flight, they were apparently more likely to nominate members of their own subgroups, perhaps because they had greater information on which to base their nominations. However, this finding is not entirely consistent with the study by Kenny and Kashy (1994), which found that friends were rated more favorably than acquaintances. In that study, mean ratings of friends were higher for positive traits, but were lower for negative traits. Because the PD traits in the present study are nearly all pejorative, we might expect that participants would be less likely to nominate people they know well for negative traits. However, the findings of Kenny and Kashy may also help to explain the smaller mean difference found in cliques compared with other subgroups. Cliques are likely to represent tight-knit groups of friends (as demonstrated by the high acquaintance scores within them), whereas the other subgroup types include both friends and acquaintances. Participants may, in fact, have been less willing to nominate their close friends as disordered, resulting in the decreased mean scores for cliques.

How does self-peer agreement within subgroups compare to between subgroups?

Having found a consistent and significant effect of mutual subgroup membership on the mean peer rating, we now turn to the question of self-peer agreement. Specifically, is self-peer agreement within subgroups higher than that between subgroups? As described above, I analyzed mixed linear models predicting peer ratings with and without the effect of self-report, and used the rater variance to calculate the self-peer correspondence. The results across all rater-target dyads, as seen in Table 14 indicate an mean effect size of .048, approximately equivalent to a self-peer correlation of .21. Table 14 also summarizes the findings across all diagnostic categories and subgroup types, demonstrating that the mean self-peer correspondence within subgroups is .05, whereas across subgroups it is .03. Although the effect sizes are small, they indicate that self report explains 56% more of the rater variance within subgroups than between subgroups.

These findings are consistent with previous studies of self-peer agreement as a function of acquaintance. Numerous cross-sectional studies have found that self-peer agreement is higher for friends than strangers (Funder & Colvin, 1988), acquaintances than strangers (Colvin & Funder, 1991), and friends than acquaintance (Kenny & Kashy, 1994). In addition, many longitudinal studies (Paulhus & Bruce, 1992; Paulhus & Reynolds, 1995; Paunonen, 1989) have found that as acquaintance increases over time, so too does self-peer agreement.

However, in the present study, acquaintance does not appear to be the only determinant in self-peer agreement. Examining agreement by subgroup type, as seen in Tables 15 through 18, it is apparent that improvement in agreement varied widely as a function of subgroup type. For example, although large improvements were seen in some within-clique and within-k-plex analyses, particularly for histrionic PD, between-group agreement was frequently higher than within-group, particularly for Cluster A disorders.

This is contrary to what we might expect if agreement were linearly related to acquaintance. As seen in Table 7, the average acquaintanceship within cliques and k-plexes was considerably higher than that within other types of subgroups. However, despite the increased acquaintanceship, cliques and k-plexes did not seem to reliably improve self-peer agreement within them.

One reason for the inconsistent effect of cliques and k-plexes may be the restrictive nature of these measures. Cliques result from very specific circumstances, in which all members of the subgroup mutually know one another very well. An individual may be excluded from a clique even if he or she knows all but one of the other members quite well (Wasserman and Faust, 1994). (K-plexes are less restrictive than cliques, but only slightly.) Because these methods are so specific, they have a tendency to exclude individuals capriciously. By excluding raters who are otherwise very similar to other subgroup members, the within-subgroup agreement is not maximized as much as might be expected, and the between-subgroup agreement is inflated.

A second explanation comes from cultural consensus theory (e.g., Romney & Weller, 1984; Romney, Weller, & Batchelder, 1986). Cultural consensus theory proposes that a rater's reliability (that is, his or her agreement with others in the group) predicts his or her accuracy. Accuracy, in terms of consensus analysis, refers to an individual's agreement with an underlying consensual "truth." This definition of accuracy is similar to the way in which Cronbach's Alpha determines how well a test item measures a construct by calculating its agreement with other similar items (Romney & Weller). Consensual analysis studies of peer perceptions of personal traits (Iannucci & Romney, 1994;

Webster, Iannucci, & Romney, 2002) have found that peers with the highest inter-rater reliability were also the most accurate relative to objective criteria.

Consensus analysis is particularly applicable to the present study, in which generalizability theory was used to model peer ratings. We postulated that a given rater's rating of a given target was a function of the target's average deviation from the grand mean, the rater's average deviation from the grand mean, and idiosyncratic factors between the target and rater. By this definition, the "true" peer rating of a target within a flight would be the rating by peers if there were no rater variance or idiosyncratic/error variance. (Note that "truth" in this sense does not imply a Platonic truth about the target, merely the consensual truth as peers see it.)

Previous studies of personality disorders have shown that peers, particularly those well-acquainted with the target, can achieve low to moderate self-peer agreement (e.g., Klonsky et al., 2002). However, no study of either normal or disordered personality has found anything close to perfect self-peer agreement, suggesting that there are, due to various factors, real differences in perceptions by self and peers (e.g., Kenny, 1994; Clifton et al., in press). In other words, there is an upper limit to self and peer agreement. Cultural consensus theory would predict that this upper limit for any group is represented by the agreement between self report and the consensual truth across all peer raters. Further, as with the generalized model described above, the self-peer agreement of any individual rater is a function of this upper limit, and the rater's deviation from the consensual truth.

Therefore, according to cultural consensus theory, self-peer agreement is directly related to a rater's reliability with the rest of the group. As seen in Tables 18 through 21, and as will be discussed at length in the next section, raters who shared a mutual subgroup with a target had considerably higher inter-rater consensus than did those who did not share a subgroup. Consensus was increased within groups, regardless of the type of subgroup studied. So, with some small differences, dividing a flight into two large clusters or factions provided the same benefits to consensus as identifying numerous small cliques, k-plexes, or exhaustive partitions. However, the larger group size for two-partition factions and clusters, relative to the other subgroup types, suggests that dividing flights into two partitions will exclude fewer reliable raters from the subgroup, therefore increasing self-peer agreement within-group relative to between-group.

To summarize, agreement between self and peers was higher within cohesive subgroups than between them. The most reliable increases in self-peer agreement, relative to between-group analyses, came from partitioning networks into two large groups, by either cluster analysis or the faction procedure. Utilizing smaller subgroups, such as cliques or k-plexes, or a greater number of partitions, particularly clusters, frequently resulted in decreased improvement relative to the between-group agreement. In order to maximize self-peer agreement, therefore, partitioning groups into two factions or clusters appears to provide the most consistent improvements.

How does rater consensus within subgroups compare to between subgroups?

The final question of interest investigates how rater-rater consensus is affected by social network subgroups. Specifically, do ratings within a cohesive subgroup agree more than those between subgroups? As described above, for each diagnostic category the rater variance and residual variance of compact mixed linear models were used to calculate the intraclass correlation of raters. These results, across all participants, are summarized in Table 18, and described by subgroup in Tables 19 through 21. Consistent with prior social network research (e.g., Carley, 1986; Freeman et al., 1987), consensus was considerably higher within subgroups than between subgroups.

Interestingly, for some diagnostic categories, not only the ratings made within subgroups, but the ratings made *between* subgroups were more reliable than the ratings across the entire group as a whole. A possible explanation for this finding may be that the mean ratings between subgroups are significantly lower overall, as seen in Tables 8 through 11. Lower mean ratings suggest that raters are nominating fewer individuals outside of their own subgroups, which inflates estimates of reliability (they are tacitly "agreeing" by not making ratings). Conversely, the fact that both within-group mean ratings and within-group reliability are significantly higher suggests that, within-subgroup, raters are agreeing on whom to nominate, rather than just whom *not* to nominate.

There are several explanations for the increased reliability within subgroups. These explanations are not necessarily competing theories, but rather it is likely that some or all act together to improve consensus. First and foremost, the increased consensus is likely associated with the increased acquaintance within subgroups relative to between subgroups. Within subgroups, raters are more highly acquainted with the targets, which numerous studies have demonstrated results in increased consensus (e.g., Kenny et al., 1994; Funder & Colvin, 1988; Kenny & Kashy, 1994). However, reliability was increased equally in both all types of subgroups, even though mean levels of acquaintance varied widely. It seems likely, therefore, that other factors also play a role in the increased consensus. Kenny's Weighted Average Model (1994) provides a comprehensive model of consensus from which to base our examination. As noted earlier, the WAM posits that consensus is a function of numerous factors. This includes rater-target acquaintance, but also rater-rater factors such as overlap, communication, and similar meaning systems. Kenny describes a set of curvilinear relationships among these factors, mediating acquaintance effects and leading to an overall consensus level. The interactions among these relationships are quite complex, and cannot be adequately modeled based on the data collected for this study. However, the cohesive clusters analyzed here can be seen as a surrogate measure for each of these factors, helping to explain the increased consensus within subgroups.

Overlap is an estimate of the amount of rater-target acquaintance that raters share. That is, if acquaintance is a measure of how many of the target's behaviors a rater has observed, overlap is a measure of the number of these behaviors two raters observed in common (Kenny, 1994). Consistent with the present study, raters within a cohesive subgroup are likely to be exposed to the same sort of information (i.e., overlap), leading them to make similar judgments. For example, Freeman and colleagues (1987) found that participants within cohesive subgroups were exposed to the same sorts of information about other members of the network, leading to similar cognitive biases. As a result, they had high consensus about other members of the network, even when these judgments were wrong.

Communication is an estimate of raters' discussion of their impressions of targets. Raters who discuss their ratings amongst themselves are likely to demonstrate higher consensus (Kenny, 1994). Within social network analysis, raters who belong to a cohesive subgroup are very often connected with one another, particularly in the more restrictive subgroups like cliques. The theoretical tradition which Pattison (1994) describes as "Information Bias" considers network connections as conduits for information, which would lead to greater communication among raters. For example, Carley (1986) noted that members of a network gained information based on their pattern of connections with others. As a result, as in the present study, Carley found that consensus was enhanced within cohesive subgroups.

Finally, similar meaning systems suggests that, in order for raters to have high consensus, they must interpret targets' behavior in the same way. Although this concept is more difficult to abstract from social network connections, there is some evidence that individuals within cohesive subgroups may have more similar meaning systems than those between subgroups. This theory of structural balance is supported by a long tradition in both social psychology (e.g., Heider, 1958; Festinger, 1957; Newcomb, 1968, et al.) and social network analysis (e.g., Johnsen, 1986). In addition, social psychological studies of group polarization (e.g., Moscovici and Zavalloni, 1969; Myers and Bishop, 1971) also indicate that small groups of individuals with similar attitudes can become polarized, such that the strength of the attitudes become enhanced for all group members.

It therefore seems likely that the increased consensus found within subgroups is at least partly due to the effects of increased similar meaning systems among group members.

Three examples may be useful in understanding how social network analysis can represent these three factors of Kenny's WAM model (overlap, communication, and similar meaning systems) to yield results similar to those of the present study. Consider Figure 1, which depicts a non-directed dichotomous network. (For simplicity, assume that if a connection is present, the connected individuals have high levels of acquaintance, whereas an absent tie indicates a low level of acquaintance.)

First, note the clique consisting of nodes 3, 4, 5, and 6. If individuals 3 and 4 are judging the personality of person 5, each has a high level of acquaintance with person 5. The fact that all members of the clique are well-acquainted with one another (the definition of a clique), makes it likely that all four clique members have spent time together. Therefore, persons 3 and 4 are likely to have observed person 5's behavior at the same time, increasing the level of overlap in their observations. In addition, because persons 3 and 4 are highly acquainted with each other, it may be more likely that they have discussed their opinions of person 5, increasing the influence of communication in their ratings. And, as members of a common social group, their interpretation of others' behavior may tend to be similar, or at least not radically different. Because persons 3 and 4 have high levels of acquaintance, overlap, communication, and similar meaning systems, we would therefore expect to find high levels of consensus between their ratings.

Next, consider persons 12 and 14 making ratings of person 11. As in the previous example, both raters are well acquainted with the target, but are not well acquainted with one another. In this case, however, persons 12 and 14 have many friends in common (persons 13, 15, and 16). Partitioning the network into factions or clusters groups all of these individuals together. Further, they are members of a mutual 2-plex, in which all members (except persons 12 and 14) are well-acquainted with one another. Therefore, although persons 12 and 14 are not directly connected, it may be more likely that they have been around person 11 at the same time, which may increase their level of overlap. Although they are not connected with one another, they may have talked about person 11 with others in their mutual subgroup. This may serve to increase their indirect communication about person 11 (through intermediaries) to a level greater than if they did not share a subgroup. As previously noted, as members of the same group, structural balance may lead them to have similar opinions. We would therefore expect good consensus among these two raters as well, despite the fact that they are not strongly connected with one another.

Finally, consider persons 5 and 12 making ratings of person 11. Both raters are acquainted with the target. They also do not share any social groups in common, making it more likely that they have observed person 11's behavior at different times and in different situations (i.e., decreased overlap).In addition, neither rater is well acquainted with the other, nor do they have any mutual friends, making it unlikely that they have discussed their opinions of person 11 either directly or indirectly (i.e., low communication). Lastly, because they are not members of a mutual subgroup, we have no way of knowing how similar their meaning systems may be. We might therefore expect to find lower levels of consensus in their ratings than in either of the previous examples.

These three examples are hypothetical, but provide a way of applying social network analysis to aspects of Kenny's model, by acting as a surrogate measure of acquaintance, overlap, communication, and similar meaning systems. Taken as a whole, these factors help to explain the enhanced consensus found in the present study.

To summarize, comparing within- and between-subgroup rater consensus found improved rater reliability within subgroups, regardless of the subgroup type used. The magnitude of this improvement varied slightly by method of subgroup, with maximal partition factions providing the greatest improvement, and maximal partition clusters providing the least. However, in all cases, within-subgroup consensus was higher than between-subgroup, and higher than that for the entire group as a whole.

Limitations and Future Research

Several limitations of the present study should be noted. First, it is unclear how well these findings will generalize to the larger population. The participants in this study were military recruits at the end of six weeks of training. Although the participants were representative of the general populace in many ways, their social situation was unique. Much more of their time and dealings with others were constrained by the requirements of training. For instance, the results of this study indicate that cohesive subgroups were strongly related to the gender of participants. As discussed earlier, this result may be due in part to the structure of being housed in separate barracks during military training. Further research to generalize these findings should attempt to extricate the effects of the military environment from subgroup formation. It may be that, in less constrained environments in which individuals are more free to choose their circles of friends, subgroups would be less influenced by gender and more influenced by personality characteristics.

That said, the constraints of military life should not be interpreted as limiting the particular findings of this study. Although the participants' social environment was affected by some unusual restrictions, it is likely that factors associated with subgroup formation in any type of social network will be difficult to generalize. The social forces at work in a college dorm are very different from those in a workplace, which are in turn different from those at a country club. More important to the findings presented here are the implications for self-peer agreement and rater-rater consensus associated with mutual subgroup membership.

A second area for improvement on the current study is an enhanced measurement of the social networks themselves. Because the data for this study was collected primarily for an investigation of self- and peer-perceptions of personality pathology, a complete assessment of social relations was limited by time and feasibility. Acquaintance was assessed based on a single question about how well the rater knew each target. In a more complete social network study, acquaintance might be assessed in alternate ways. Rather than simply asking "how well do you know," participants might be asked more specific questions regarding friendship, advice seeking, amount of time spent with, and other aspects of acquaintance. Rankings of acquaintance, rather than categorical ratings could provide more specific information about connections within the network. Longitudinal analyses, examining changes in both acquaintance and personality ratings, would be very helpful in understanding network associations with ratings. In addition, particularly given the focus of the present research, the network structure might be assessed in ways other than self report. Asking participants to identify friendships between other dyads, or observations by an outside party might yield a different picture of the network than that derived from self-report data alone (e.g., Bernard et al., 1980).

Implications and Conclusions

Several conclusions may be drawn from the results of the present study, leading to potential practical applications of the findings. As discussed above, it is unclear how well these specific findings will generalize to other populations, given the unique social structure studied here. However, I believe that although the specific predictors of cohesive subgroups may be less applicable in other settings, the underlying theories described here may provide real benefits to researchers collecting peer report data.

First, consistent with prior research (e.g., Klonsky et al., 2002), agreement between self and peers was low. This was true both for aggregated peer data, and when calculated for individual rater-target dyads. Correlations for diagnostic scales within a domain (i.e., self-report or peer-report) were much stronger than those of scales across domains (i.e., self-peer agreement). However, in most cases self-peer agreement was highest for corresponding scales than for non-corresponding scales, suggesting that self and peers are at least rating a similar construct, even if their interpretations of it differ (e.g., Clifton et al., in press).

Second, several predictors for network position were identified. The subgroups in this study appear to be strongly related to both gender and race, although race was not nearly as assortative as gender. In addition, numerous pathological personality traits were significant predictors of network position. In general these associations were consistent with the DSM-IV descriptions of the personality disorders, such that Cluster B scales were associated with increased social connections and a more central position in the network, whereas Cluster A and Cluster C scales were negatively associated with these characteristics. Although effect sizes were small, aggregated peer report was generally a better predictor of network position than self report, providing additional evidence for the validity of peer reports of personality disorder.

Third, ratings made of individuals within one's own subgroup resulted in higher self-peer agreement than ratings of individuals outside of the subgroup. The magnitude of this effect varied by diagnostic category and method of subgroup determination. However, the improvements from dividing networks into two partitions based on faction or cluster analysis were more consistent and, in most cases, larger than those obtained from a greater number of small subgroups. Dividing a network into two partitions seems to be adequate to increase self-peer agreement, whereas increasing the number of partitions may exclude reliable raters from the analysis.

Finally, ratings made within a subgroup demonstrated higher peer-peer reliability than those made between subgroups. Although improvements in reliability also varied by diagnostic category and subgrouping method, improvements were more consistent across methods than those for self-peer agreement. This increased consensus can be understood as some combination of increased acquaintance, overlap, communication, and similarity in meaning systems within subgroups as compared to between.

These findings are directly applicable to future studies of peer perceptions of personality. Relatively few studies of personality pathology assess peer perceptions, relying instead on self report. Both the present study and previous research (e.g., Klonsky et al., 2002) have found that peer perspectives of personality differ substantially from those of the self. Further, the present study is consistent with earlier findings (e.g., Kolar, Funder, and Colvin, 1996) suggesting that, in some cases, peer reports of personality may predict objective behavior (in this case, network position) better than self report. Taken together, these findings underscore the importance of obtaining information from peers whenever possible.

The vast majority of research on peer reports of personality pathology has relied on only one or two selected informants for information. As noted previously, this practice both limits the reliability of the data and may suffer from a selection bias. Obtaining data from larger groups of peers is preferable, but also methodologically difficult. One of the complications of obtaining data from large peer groups is that not all peers know one another equally well. Although obtaining reports from multiple peers increases the maximum possible reliability of the data, it by no means ensures reliable ratings. The present study indicates that by identifying cohesive subgroups within networks of raters, rater reliability was increased by 40 to 50%. The increase in mean ratings and overall variance suggests that peers may be more reliable within subgroups because they are both making more nominations within their subgroups and agreeing on them more. Increasing rater reliability by removing unreliable raters decreases the “noise” in the analysis, and increases the upper limits of validity.

That the validity of the peer data was improved is implied, but not assured, by the increased self-peer agreement within subgroups relative to that between subgroups. Obviously, greater self-peer agreement does not imply that the data is any closer to the hypothetical truth about an individual. Peers within an individual’s subgroup might simply subscribe to the same false beliefs as the individual. The fact that self and peer reports converge within subgroups is promising. However, without a gold standard measure of an individual’s personality, it is unclear whether subgroups predict improved accuracy, or merely improved agreement.

Based on the findings of the present study, I would encourage the incorporation of network analysis techniques into the assessment of normal and maladaptive personality. Obviously a full social network analysis is time consuming and beyond the scope of most personality research. However, when collecting peer reports, particularly in round-robin designs, it is relatively simple to also ask about the level of acquaintance for each participant. The difficulty comes in incorporating both rater-rater and rater-target acquaintance levels into subsequent data analyses. Social network analysis techniques such as the factions routine (in UCINET 6) or cluster analysis (which most researchers have access to without need of a specialized network analysis computer program) can simplify this process by identifying cohesive groups within the acquaintance network. Participants can then be assigned a single categorical variable, corresponding to their subgroup membership, which is easily incorporated into analysis. The results of the present study suggest that partitioning a network into two groups provides as much benefit, if not more, to reliability and self-peer agreement compared with identifying very specific cohesive subgroups like cliques and k-plexes. Researchers collecting whole-network peer report data might consider partitioning data into two cohesive subgroups as a relatively simple and efficient way to improve reliability and accuracy. Incorporating social network techniques into personality research may be a small but important step toward greater understanding of the relationship between self and peer perceptions of personality pathology.

References

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55*, 387-395.

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster Analysis*. Newbury Park, CA: Sage Publications.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders (4th ed.)*. Washington, DC: Author.

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.

Bales, R. F., Cohen, S. P., & Williamson, S. A. *SYMLOG, a system for the multiple level observation of groups*. New York: Free Press.

Batagelj, V., & Mrvar, A. (2001). Pajek program for large network analysis (Version 0.79). [Computer software]. Ljubljana, Slovenia.

Bernard, H. R., & Killworth, P. D. (1977). Informant accuracy in social network data: II. *Human Communication Research, 4*(1), 3-18.

Bernard, H. R., Killworth, P. D., & Sailer, L. (1982). Informant accuracy in social-network data V: An experimental attempt to predict actual communication from recall data. *Social Science Research, 11*, 30-66.

Bernard, H. R., Killworth, P. D., Kronenfeld, D., & Sailer, L. (1984). The problem of informant accuracy: The validity of retrospective data. *Annual Review of Anthropology, 13*, 495-517.

Bingham, W. V. (1939). Halo, invalid and valid. *The Journal of Applied Psychology, 23*, 221-228.

Bonacich, P. (1987). Power and centrality: a family of measures. *American Journal of Sociology, 92*, 1170-1182.

Borgatti, S. P., & Everett, M. G. (1989). The class of all regular equivalences: Algebraic structure and computation. *Social Networks, 11*, 159-172.

Borgatti, S. P. (1997). Structural holes: Unpacking Burt's redundancy measures. *Connections, 20*, 35-38.

Borgatti, S.P., Everett, M.G., & Freeman, L.C. (2002). Ucinet 6 for Windows. [Computer software and manual]. Harvard: Analytic Technologies.

Brass, D. J. (1995). A social network perspective on human resources management. *Research in Personnel and Human Resources Management, 13*, 39-79.

Breiger, R. L., & Ennis, J. G. (1979). Personae and social roles: The network structure of personality types in small groups. *Social Psychology Quarterly, 42*(3), 262-270.

Bron, C., & Kerbosch, J. (1973). Finding all cliques of an undirected graph. *Communications of the ACM, 16*, 575-577.

Brutus, S., Fleenor, J. W., & McCauley, C. D. (1999). Demographic and personality predictors of congruence in multi-source ratings. *Journal of Management Development, 18*, 417-435.

Burt, R. S., & Bittner, W. M. (1981). A note on inferences regarding network subgroups. *Social Networks, 3*, 71-88.

Burt, R. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology, 92*, 1287-1335.

Burt, R. S., Jannotta, J. E., & Mahoney, J. T. (1998). Personality correlates of structural holes. *Social Networks, 20*(1), 63-87.

Carley, K. M. (1986). An approach for relating social structure to cognitive structure. *Journal of Mathematical Sociology, 12*, 137-189.

Casciaro, T. (1998). Seeing things clearly: Social structure, personality, and accuracy in social network perception. *Social Networks, 20*(4), 331-351.

Casciaro, T., Carley, K. M., & Krackhardt, D. (1999). Positive affectivity and accuracy in social network perception. *Motivation & Emotion, 23*(4), 285-306.

Chu, S.-C., & Roddick, J. F. (2000). A clustering algorithm using the tabu search approach with simulated annealing. In N. Ebecken & C. A. Brebbia (Eds.), *Data Mining II/ International Conference on Data Mining* (pp. 515-523). Cambridge, England: WIT Press.

Clark, L. A. (1993). *Schedule for Nonadaptive and Adaptive Personality (SNAP)*. Minneapolis, MN: University of Minnesota Press.

Clifton, A., Turkheimer, E., & Oltmanns, T. F. (in press). Contrasting perspectives on personality problems: Descriptions from the self and others. *Personality and Individual Differences.*

Colvin, C. R., & Funder, D. C. (1991). Predicting personality and behavior: A boundary on the acquaintanceship effect. Journal of Personality & Social Psychology, 60, 884-894.

Cronbach, L. (1955). Processes affecting scores on "understanding of others" and "assumed similarity.". *Psychological Bulletin, 52*, 177-193.

Cronbach, L. J., & Gleser, G. (1953). Assessing similarity between profiles. *Psychological Bulletin, 50*, 456-473.

Cronbach, L., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability of scores and profiles*. New York: Wiley.

de Amorim, S. G., Barthelemy, J.-P., & Ribeiro, C. C. (1992). Clustering and Clique Partitioning: Simulated Annealing and Tabu Search Approaches. *Journal of Classification, 9*, 17-41.

Dean, J. W., & Brass, D. J. (1985). Social interaction and the perception of job characteristics in an organization. *Human Relations, 38*, 571-582.

Doreian, P. (1979). *Mathematics and the study of social relations*. London: Weidenfeld & Nicholson.

Dowson, J. H. (1992). DSM-III--R narcissistic personality disorder evaluated by patients' and informants' self-report questionnaires: Relationships with other personality disorders and a sense of entitlement as an indicator of narcissism. *Comprehensive Psychiatry, 33*(6), 397-406.

Everett, M. G. (1985). Role similarity and complexity in social networks. *Social Networks, 7*, 353-359.

Everett, M. G. (n.d.). *4 cohesive subgroups*, [webpage]. Available: www.analytictech.com/networks/chapter_4_cohesive_subgroups.htm.

Ferro, T., & Klein, D. N. (1997). Family history assessment of personality disorders: I. Concordance with direct interview and between pairs of informants. *Journal of Personality Disorders, 11*(2), 123-136.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Field, S. L., & Carter, W. C. (2002). Detecting measurement bias in respondent reports of personal networks. *Social Networks, 24*, 365-383.

Frank, K. A. (1995). Identifying cohesive subgroups. *Social Networks, 17*, 27-56.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*, 35-41.

Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks, 1*, 215-239.

Freeman, L. C., & Romney, A. K. (1987). Words, deeds, and social structure: A preliminary study of the reliability of informants. *Human Organization, 46*(4), 330-334.

Freeman, L. C., Romney, A. K., & Freeman, S. C. (1987). Cognitive structure and informant accuracy. *American Anthropologist, 89*(2), 310-325.

Friedkin, N. E. (1991). Theoretical foundations for centrality measures. *American Journal of Sociology, 96*, 1478-1504.

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality & Social Psychology, 55*, 149-158.

Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science, 2*(1), 31-39.

Funder, D. C., & West, S. G. (1993). Consensus, self-other agreement, and accuracy in personality judgment: An introduction. *Journal of Personality, 61*, 457-476.

Gage, N. L., & Cronbach, L. (1955). Conceptual and methodological problems in interpersonal perception. *Psychological Review, 62*, 411-422.

Galaskiewicz, J., & Burt, R. S. (1991). Interorganization contagion in corporate philanthropy. *Administrative science quarterly, 36*, 88-105.

Glover, F. (1989). Tabu search- Part I. *ORSA Journal on Computing, 1*(3), 190-206.

Glover, F. (1990). Tabu search- Part II. *ORSA Journal on Computing, 2*(1), 4-32.

Glover, F., & Laguna, M. (1997). *Tabu search*. Boston: Kluwer Academic Publishers.

Gormly, J., & Edelberg, W. (1974). Validity in personality trait attribution. *American Psychologist, 29*, 189-193.

Gould, R.V. (1987). Measures of betweenness in non-symmetric networks. *Social Networks, 9*, 277-282.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Iannucci, A. L., & Romney, A. K. (1994). Consensus in the judgment of personality traits among friends and acquaintances. *Journal of Quantitative Anthropology, 4*, 279-295.

Ibarra, H., & Andrews, S. B. (1993). Power, social influence, and sense making: Effects of network centrality and proximity on employee perceptions. *Administrative Science Quarterly, 38*(2), 277-303.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61*(4), 521-551.

John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality & Social Psychology, 66*(1), 206-219.

Johnsen, E. C. (1986). Structure and process: Agreement models for friendship formation. *Social Networks, 8*, 257-306.

Kane, J. S., & Lawler, E. E., III. (1978). Methods of peer assessment. *Psychological Bulletin, 85*, 555-586.

Kanfer, A., & Tanaka, J. S. (1993). Unraveling the web of personality judgments: The influence of social networks on personality assessment. *Journal of Personality, 61*(4), 711-738.

Kashy, D. A., & Kenny, D. A. (1990). Do you know whom you were with a week ago Friday? A re-analysis of the Bernard, Killworth, and Sailer studies. *Social Psychology Quarterly, 53*(1), 55-61.

Kenny, D. A., & LaVoie, L. (1984). The Social Relations Model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 142-182). Orlando, FL: Academic Press.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, NY: The Guilford Press.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin, 116*, 245-258.

Kenny, D. A., & Kashy, D. A. (1994). Enhanced co-orientation in the perception of friends: A social relations analysis. *Journal of Personality & Social Psychology, 67*, 1024-1033.

Kenny, D. A. (1996). The design and analysis of social-interaction research. *Annual Review of Psychology, 4, 59-86.*

Kenny, D. A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D. A. (2002). The statistical analysis of data from small groups. *Journal of Personality & Social Psychology, 83*(1), 126-137.

Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist, 43*(1), 23-34.

Kilduff, M. (1990). The interpersonal structure of decision making: A social comparison approach to organizational choice. *Organizational behavior and human decision processes, 47*, 270-288.

Kilduff, M. (1992). The friendship network as a decision-making resource: Dispositional moderators of social influences on organizational choice. *Journal of Personality and Social Psychology, 62*(1), 168-180.

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.

Krackhardt, D., & Kilduff, M. (1999). Whether close or far: Social distance effects on perceived balance in friendship networks. *Journal of Personality & Social Psychology, 76*, 770-782.

Lawler, E. L., Lenstra, J. K., Kan, G. R., & Shmoys, D. B. (Eds.). (1990). *The traveling salesman problem: A guided tour of combinatorial optimization*. Chichester, New York: Wiley.

Lenzenweger, M. F., Loranger, A. W., 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.

Lorrain, F., & White, H. C. (1971). The structural equivalence of individuals in social networks. *Journal of Mathematical Sociology, 1*, 49-80.

Malek, M., Guruswamy, M., Owens, H., & Pandya, M. (1989). Serial and parallel search techniques for the traveling salesman problem. *Annals of operations research, 21*, 59-84.

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.*

Malloy, T. E., Albright, L., Kenny, D. A., Agatstein, F., & et al. (1997). Interpersonal perception and metaperception in nonoverlapping social groups. *Journal of Personality & Social Psychology, 72*(2), 390-398.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.

Mehra, A., Kilduff, M., & Brass, D. J. (2001). The social networks of high and low self-monitors: Implications for workplace performance. *Administrative Science Quarterly, 46*(1), 121-146.

Michaelson, A., & Contractor, N. S. (1992). Structural position and perceived similarity. *Social Psychology Quarterly, 55*(3), 300-310.

Mokken, R. J. (1979). Cliques, clubs, and clans. *Quality and Quantity, 13*, 161-173.

Moody, J. (2001). Peer influence groups: Identifying dense clusters in large networks. *Social Networks, 23*(4), 261-283.

Moreno, J. L. (1934). *Who Shall Survive?* Washington, DC: Nervous and Mental Disease Publishing Company.

Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology, 12*, 125-135.

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.

Myers, D. G., & Bishop, G. D. (1971). The enhancement of dominant attitudes in group discussion. *Journal of Personality and Social Psychology, 20*, 386-391.

Newcomb, T. M. (1968). Interpersonal balance. In R. B. Abelson (Ed.), *Theories of cognitive consistency: A sourcebook* (pp. 28-51). Chicago: Rand McNally.

Nieminen, U. J. (1973). On the centrality in a directed graph. *Social Science Research, 2*, 371-378.

Oltmanns, T. F., Turkheimer, E., & Strauss, M. E. (1998). Peer assessment of personality traits and pathology in female college students. *Assessment, 5*, 53-65.

Pan, J. S., Chu, S.-C., & Lu, Z. M. (2000). Cluster generation using tabu search based maximum descent algorithm. In N. Ebecken & C. A. Brebbia (Eds.), *Data Mining II/ International Conference on Data Mining* (pp. 525-534). Cambridge, England: WIT Press.

Park, B., & Judd, C. M. (1989). Agreement on initial impressions: Differences due to perceivers, trait dimensions, and target behaviors. *Journal of Personality & Social Psychology, 56*, 493-505.

Park, B., DeKay, M. L., & Kraus, S. (1994). Aggregating social behavior into person models: Perceiver-induced consistency. *Journal of Personality & Social Psychology, 66*, 437-459.

Pattison, P. (1994). Social cognition in context: Some applications of social network analysis. In S. Wasserman & J. Galaskiewicz (Eds.), *Advances in social network analysis: Research in the social and behavioral sciences* (pp. 79-109). Thousand Oaks, CA: Sage.

Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63*(5), 816-824.

Paulhus, D. L., & Reynolds, S. (1995). Enhancing target variance in personality impressions: Highlighting the person in person perception. *Journal of Personality and Social Psychology, 69*(6), 1233-1242.

Paunonen, S. V. (1989). Consensus in personality judgments: Moderating effects of target-rater acquaintanceship and behavior observability. *Journal of Personality & Social Psychology, 56*, 823-833.

Popielarz, P. A. (2000). *Connecting structure and content: Density and racial/ethnic homogeneity in personal networks.* Paper presented at the 20th International Sunbelt Social Network Conference, Vancouver, BC, Canada.

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.

Riso, L. P., Klein, D. N., Anderson, R. L., Ouimette, P. C., & et al. (1994). Concordance between patients and informants on the personality disorder examination. *American Journal of Psychiatry, 151*(4), 568-573.

Romney, A. K., & Faust, K. (1982). Predicting the structure of a communications network from recalled data. *Social Networks, 4*(4), 285-304.

Romney, A. K., & Weller, S. C. (1984). Predicting informant accuracy from patterns of recall among individuals. *Social Networks, 6*(1), 59-77.

Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist, 88*, 313-338.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika, 31*, 581-603.

Salhi, S. (1998). Heuristic search methods. In G. A. Marcoulides (Ed.), *Modern methods for business research: Methodology for business and management* (pp. 147-175).

SAS Institute Inc. (1999). SAS OnlineDoc, Version 8. [Computer manual].Cary, NC: SAS Institute Inc. Retrieved October 23, 2002, from http://v8doc.sas.com/sashtml/

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.

Scott, J. P. (2000). *Social Network Analysis: A Handbook* (2nd ed.). London: Sage Publications.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Thousand Oaks, CA: Sage Publications, Inc.

Shrout, P. E., & Fliess, J. L. (1979). Intraclass correlation: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Shrout, P. E. (1993). Analyzing consensus in personality judgments: A variance components approach. *Journal of Personality, 61*(4), 769-788.

Shrout, P. E. (1995). Measuring the degree of consensus in personality judgments. In P. E. Shrout & S. T. Fiske (Eds.), *Personality research, methods, and theory: A festschrift honoring Donald W. Fiske* (pp. 79-92). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24*, 323-355.

Snijders, T. A. B., & Bosker, R. J. (1994). Modeled variance in two-level models. Sociological Methods and Research, 22, 342-363.

Strauss, J. P., Barrick, M. R., & Connerley, M. L. (2001). An investigation of personality similarity effects (relational and perceived) on peer and supervisor ratings and the role of familiarity and liking. *Journal of Occupational & Organizational Psychology, 74*, 637-657.

Taft, R. (1959). Multiple methods of personality assessment. *Psychological Bulletin, 56*, 333-352.

Tajfel, H. (Ed.). (1978). *Differentiation between social groups*. New York: Academic Press.

Terry, R., & Coie, J. D. (1991). A comparison of methods for defining sociometric status among children. *Developmental Psychology, 27*, 867-880.

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.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications*. New York: Cambridge University Press.

Webster, C. M., Iannucci, A. L., & Romney, A. K. (2002). Consensus analysis for the measurement and validation of personality traits. *Field Methods, 14*, 46-64.

Weissman, M. 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.

Zimmerman, M., Pfohl, B., Coryell, W., Stangl, D., & Corenthal, C. (1988). Diagnosing personality disorder in depressed patients: A comparison of patient and informant interviews. *Archives of General Psychiatry, 45*(8), 733-73

Appendix

Items comprising the Peer Inventory of Personality Disorder.

The 81 items based on lay translations of the DSM-IV criteria are indicated by the corresponding diagnostic category in parentheses.

1. Please select the people who are your close friends
2. Please select the people who you don't know at all
3. Is sympathetic and kind to others
4. Has a good sense of humor
5. Is trustworthy and reliable
6. Is articulate and persuasive in a discussion
7. Has a cheerful and optimistic outlook on life
8. Acts as a leader
9. Seems to be quite independent
10. Prefers to do things alone (Schizoid)
11. Is superstitious or believes in mind-reading (Schizotypal)
12. Is likely to pursue a task diligently until it is completed
13. Seems to feel empty inside (Borderline)
14. Spends too much time thinking about gaining unlimited success, power, or love (Narcissistic)
15. Is reserved or shy when meeting new people because he/she feels inadequate (not as good as other people) (Avoidant)
16. Needs to have other people take care of him/her (Dependent)
17. Needs to do such a perfect job that nothing ever gets finished (OCPD)
18. Is assertive in interactions with others
19. Is quite feminine; acts in a way you'd expect a female to act
20. Is not interested in close relationships
21. Is odd or peculiar in behavior or appearance (Schizotypal)

22. Lies to people, or cons people (Antisocial)
23. Lacks a fundamental sense of who he/she is (Borderline)
24. Has shallow emotions that change rapidly (Histrionic)
25. Needs other people to admire him/her (Narcissistic)
26. Worries that other people will criticize or reject him/her (Avoidant)
27. Is unrealistically and persistently afraid of being left alone to care for him/herself (Dependent)
28. Is very rigid and stubborn (OCPD)
29. Is compassionate and responds with concern when presented with others' problems
30. Has frequent doubts about the loyalty of friends; doesn't trust anyone (Paranoid)
31. Does not want to be close physically or emotionally to other people, even his/her family (Schizoid)
32. Thinks other people are talking about or looking at him/her when they aren't (Schizotypal)
33. Does things without thinking; doesn't plan ahead (Antisocial)
34. Has strong mood swings in response to events; frequent periods of feeling intense sadness, irritation, or anxiety (Borderline)
35. Talks in a vague way that lacks detail and is hard to understand (Histrionic)
36. Is stuck up or 'high and mighty' (Narcissistic)
37. Thinks that he/she is clumsy, unattractive, or inferior to other people (Avoidant)
38. After he/she breaks up with a girlfriend/boyfriend, he/she quickly finds someone else to take care of him/her (Dependent)
39. Is very stingy with money (OCPD)
40. Is generally agreeable and cooperative with others
41. Does not want to tell personal information to anyone because they might spread it around or use it against him/her (Paranoid)
42. Doesn't enjoy doing anything (Schizoid)
43. shows emotional responses that seem strange or 'out of sync' (Schizotypal)
44. Is irresponsible; can't be counted on to do his/her work or pay bills (Antisocial)

45. Has unstable, intense relationships with other people; often switches back and forth between loving a person and hating him/her (Borderline)
46. Repeatedly attempts (or threatens to attempt) suicide or to seriously harm him/herself (Borderline)
47. Behaves as if 'on stage', as if he/she is an actor; exaggerated expressions of emotion (Histrionic)
48. Is not concerned about other people's feelings or needs (Narcissistic)
49. Is unwilling to do new things because they might be embarrassing (Avoidant)
50. Feels helpless or uncomfortable when left alone; afraid that he/she won't be able to take care of him/herself (Dependent)
51. Needs to do everything him/herself because no one else will do them right (OCPD)
52. Is aggressive; tends to start arguments with other people
53. Is gentle with others
54. Reads hidden meanings into innocent things that people say or do; thinks people are putting him/her down or threatening him/her when they are not (Paranoid)
55. Has no close friends (other than family members) (Schizoid/Schizotypal)
56. Has an odd way of thinking, and his/her speech sometimes does not make sense (Schizotypal)
57. Gets mad easily and often gets in fights (Antisocial)
58. Is unmoved, and doesn't feel guilt, after hurting someone or stealing (Antisocial)
59. Has sudden, even violent outbursts of anger (Borderline)
60. Is easily influenced by other people (suggestible) (Histrionic)
61. Thinks other people are jealous of him/her (Narcissistic)
62. Is very controlled or inhibited with close friends because he/she is afraid people will make fun of him/her (Avoidant)
63. Doesn't like to disagree with other people because they might reject him/her (Dependent)
64. Can't throw out old things even if they are of no use to him/her (Obsessive-Compulsive)
65. Is dominant in his/her interpersonal relationships

66. Is charismatic and has leadership abilities
67. Is cold; doesn't show any feelings (Schizoid/Schizotypal)
68. Repeatedly gets in trouble with the police (Antisocial)
69. Will do almost anything to keep from being left alone (Borderline)
70. Gets paranoid or has brief periods of very strange behavior (acts crazy) in response to stress (Borderline)
71. Considers his/her relationships with other people to be closer (more intimate) than hey really are (Histrionic)
72. Feels he/she deserves special favors or treatment (Narcissistic)
73. Avoids working in teams, because he/she is afraid someone will criticize or reject him/her (Avoidant)
74. Can't make a simple decision without lots of advice from other people (Dependent)
75. Is afraid to do things by him/herself (Dependent)
76. Is much too concerned about details, rules, lists and schedules (OCPD)
77. Is sensitive to the needs of others
78. Displays much creativity or artistic talent
79. Thinks other people are attacking his/her reputation and reacts with anger, even though his/her friends do not see these attacks (Paranoid)
80. Doesn't care whether other people praise or criticize him/her (Schizoid)
81. Is nervous around other people because he/she doesn't trust them (Schizotypal)
82. Has a reckless lack of concern for safety of self or other people (Antisocial)
83. Is unhappy when he/she is not the center of attention (Histrionic)
84. Is unwilling to get involved with other people unless he/she is certain of being liked (Avoidant)
85. Goes to excessive lengths (will do almost anything) to get other people to take care of him/her (Dependent)
86. Works so much that he/she never has fun and has no friends (OCPD)
87. Is generally a tender person
88. Is sincere and genuine
89. Thinks that people are taking advantage of, lying to, or harming him/her (Paranoid)

90. Seems to see, hear, or experience things differently from the way other people do (Schizotypal)
91. Lives a reckless lifestyle; does dangerous things without planning (Borderline)
92. Is inappropriately sexually seductive when interacting with other people (Histrionic)
93. Thinks he/she is much better than other people (without good reason) (Narcissistic)
94. Has very strict and rigid ideas about morals and ethics (OCPD)
95. Is willing to take a stand for something he/she believes in
96. Is quite masculine; acts in a way you'd expect a male to act
97. Is suspicious that his/her sexual partner might be cheating on him/her (Paranoid)
98. Uses physical appearance to draw attention to him/herself (Histrionic)
99. Thinks that he/she is special, so he/she should only hang out with other special people (Narcissistic)
100. Remains calm and copes successfully in stressful situations
101. Seems to be quite understanding
102. Is overly suspicious or paranoid (Schizotypal)
103. Takes advantage of other people with no intention of paying them back (Narcissistic)
104. Holds grudges for a long time if insulted or injured (Paranoid)
105. Is jealous of other people (Narcissistic)