



A SOCIAL NETWORK APPROACH TO PEER ASSESSMENT: IMPROVING PREDICTIVE VALIDITY

GIL LURIA AND YUVAL KALISH

Our premise is that the simple measure used in peer assessment (i.e., number of peer nominations) does not capture the complexity of social information processing and therefore has limited predictive validity. Based on indicators derived from social network analysis and social information processing theories, we suggest new measures (nominations-by-nominees and nominations-not-returned) to enhance the predictive validity of peer assessment. We then compare the validity of existing measures with ours, using a longitudinal sample of 249 soldiers, divided into 18 groups. The soldiers first assessed each other on friendly behavior and instrumental contribution to the team. More than six months later, the commanders of the 132 soldiers in the unit under review provided evaluations of their performance in regard to stress, engagement, and leadership. We found that our new, complex measures predicted performance above and beyond the traditional measure. Theoretical and applied implications are discussed.

Keywords: peer assessment, social network analysis, social information processing, egocentric network density, nonreciprocity, indegree centrality, performance, longitudinal data

Introduction

Peer assessment is central to human resource management and is linked to many other human resource management decisions and activities. It is frequently used in organizations in decisions about employee promotion (Rodgers, 1992), selection for leadership roles (Geoffrey, 1994), developing teamwork

(Willey & Gardner, 2009), promoting organizational change (Klagge, 1995), evaluating informal organizational structure (Murray, 1970), personnel selection (Colarelli & Boos, 1992), and job placement and/or termination (London & Smither, 1995).

Although supervisors have traditionally been the only source of appraisal in many organizations (and are still considered as

Correspondence to: Gil Luria, Department of Human Services, Faculty of Social Welfare & Health Sciences, University of Haifa, Mount Carmel, Haifa 31905, Israel, Phone: 972-4-8288663, Fax: 972-4-8249282, E-mail: gluria@univ.haifa.ac.il.

Human Resource Management, July–August 2013, Vol. 52, No. 4. Pp. 537–560

© 2013 Wiley Periodicals, Inc.

Published online in Wiley Online Library (wileyonlinelibrary.com).

DOI:10.1002/hrm.21541

an important source of evaluation), it has long been recognized that single-source evaluations are problematic (see, e.g., Kane & Lawler, 1979; Milliman, Zawacki, Norman, Powell, & Kirksey, 1994; Wexley & Klimoski, 1984). Today, jobs are broader and include considerable work in groups. This enables higher exposure of peers to each other in comparison to the past (Katzenbach & Smith, 1993). Furthermore, as work often involves collaboration with colleagues who are not necessarily co-located or in the same organization (Ancona & Caldwell, 1992; Reagans & Zuckerman, 2001), there are advantages to obtaining diverse, external sources of information (Allen & Cohen,

1969; Tushman & Scanlan, 1981) on an employee's performance. To address this complexity, organizations often introduce multiple appraisals by both peers and supervisors (Bettenhausen & Fedor, 1997; London & Smither, 1995; Milliman et al., 1994). This study focuses on peer assessment (or peer appraisal, these terms being used interchangeably throughout this article), defined as "having the members of a group judge the extent to which each of their fellow group members has exhibited specified traits, behaviors, or achievements" (Kane & Lawler, 1978, p. 555).

Research has indicated the importance of peer assessment, which has been found to provide information complementary to supervisor evaluation by tapping other dimensions of performance (Borman, 1974; Zammuto,

London, & Rowland, 1982). Peers' evaluations are based on many observations by multiple observers. Such assessments are stable (Smith, 1967) and good for evaluating skills that are conducive to improved performance (Yammarino & Waldman, 1993), as well as presenting accurate judgments of coworkers' behavior (Kane & Lawler, 1978; Lewin & Zwany, 1976; Murphy & Cleveland, 1991; Wexley & Klimoski, 1984). Furthermore,

predictive validity studies have shown that peer assessments accurately predict job performance (Kane & Lawler, 1978; Mayfield, 1970; Reilly & Chao, 1982) and advancement (Kraut, 1975; Roadman, 1964; Shore, Shore, & Thornton, 1992). It seems that what validates peer assessment is both extensive exposure to peers' behavior and the fact that it provides multiple-source evaluations.

Although peer appraisal is often used in organizations, very limited research has focused on the subject, so that there is a dearth of information for organizational guidance (Bettenhausen & Fedor, 1997). We found no studies of the validity or the psychometric properties of the measures for analyzing peer assessment data. Furthermore, all the studies we reviewed used only basic aggregation measures (e.g., number of nominations or average score) in their analysis of peer assessment data. Similarly to London and Smither (1995), we suggest that such measures cannot capture the complexities and multidimensionality of job performance. More importantly, we suggest that such measures are not aligned with established social information processing such as categorization (Hogg & Terry, 2000) that is known to occur when perceiving others in a group. We therefore suggest that the existing measures used to analyze peer assessment data do not utilize the full potential of the assessment data, because part of the obtained information is lost during aggregation of such complex data (Krackhardt, 1987). More importantly, we argue that the predictive validity of peer appraisal can be improved by incorporating measures derived from a social network perspective that integrate the socio-psychological processes by which individuals construct their social world.

The goal of this study is to test whether measures that capture the complexities of interpersonal ratings of social, analytical, and job performance improve the predictive validity of peer assessment data. Based on social categorization processes (Hogg & Terry, 2001), we suggest two measures that can be incorporated in analysis of peer assessment to enhance the predictive validity of peer appraisal. The article is structured as follows:

Research has indicated the importance of peer assessment, which has been found to provide information complementary to supervisor evaluation by tapping other dimensions of performance.

introduction of social information-processing processes that are known to operate when people are expected to select others in their group; description of the existing measure for peer assessment (number of peer nominations) and its limitations; description of two measures (nominations-by-nominees and nominations-not-returned) that can overcome these limitations since they better account for the hierarchies of individual social information processing.

Social Information Processing and the Prototypicality Gradient

People use concepts—mental representations of categories—to make sense of their world (Kunda, 2002). Eleanor Rosch's (Rosch, 1978; Rosch & Mervis, 1975) seminal studies suggested that people compare a natural object (e.g., a pigeon) to a prototype, an abstraction that they have in their minds (e.g., a bird prototype), and this comparison guides their expectations, behaviors, and even emotional responses to that object. Rosch's studies also show that some objects are perceived as more prototypical (i.e., as better matches to the prototype) than others. Thus, for example, a pigeon is more aligned (or prototypical) with their concept of "bird" than an ostrich.

The social information approaches suggest that, as with classification of natural objects, people construct their social world based on the information they perceive, the context in which they are embedded, and the prototypes they have (Leyens & Fiske, 2002).

Social identity theory (Tajfel & Turner, 1979) is based on social comparisons intended to confirm or establish differences among group members. In his self-categorization theory (an extension of social identity theory), Turner (1985) proposes that social categorization gives rise to depersonalization of self and others and generates social phenomena, such as social hierarchy. Thus, social categorization divides members of a group into those who are perceived as either more or less resembling a prototype (Hogg, 1992, 1993, 2001).

According to self-categorization theory (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), when making sense of the social

world, especially in cognitively complex situations, depersonalization occurs, so that social cognition is guided by prototypicality (Hogg & Terry, 2000). In complex tasks such as peer assessment, simple, clear prototypes reduce uncertainty (Brewer & Harasty, 1996; Campbell, 1958; Hamilton, Sherman, & Lickel, 1998). As Hogg and Terry (2000, p. 126, *italics added*) argue: "Within groups people are able to distinguish themselves and others in terms of how well they match the prototype. *An intragroup prototypicality gradient exists—Some people are or are perceived to be more prototypical than others,*" and even subtle differences in prototypicality have significant impact on social information processing and individual behavior (Haslam, Oakes, McGarty, Turner, & Onorato, 1995). Empirical studies have proven that the level of prototypicality influences social attraction (Hogg, Cooper-Shaw, & Holzworth, 1993; Hogg & Hains, 1998), intragroup structure (Hornsey & Hogg, 1999, 2000), and within-group rejection of members (Marques & Paez, 1994).

To summarize, social identity (Tajfel & Turner, 1979) and self-categorization theories (Turner et al., 1987) suggest that social information processing is guided by prototypicality, especially in complex tasks. When depersonalization occurs, social complexity is reduced, so that individuals assess others according to a prototypicality gradient. This gradient is hierarchical, so that some individuals are viewed as more prototypical than others. We argue that the complexities of peer assessment may involve assessments based on prototypicality, and therefore analysis of peer assessment data should incorporate measures that integrate the prototypicality gradient. Such social comparisons can be captured by exploring how pairs of peers assess each other, assuming that a more prototypical member is more likely to be selected, but not to select others; and by analyzing the level of prototypicality of group members who select

The goal of this study is to test whether measures that capture the complexities of interpersonal ratings of social, analytical, and job performance improve the predictive validity of peer assessment data.

a peer, under the assumption that a more prototypical group member is a better assessor.

We therefore present two additional measures that are more attuned to the hierarchical nature of social information processing, and thus provide better validity in regard to peer appraisal. After presenting the standard measure of peer nominations, and discussing its strengths and limitations, we propose two other measures: namely, nominations-not-returned and nominations-by-nominees.

Peer Nominations Measure

This measure taps into the prototypicality gradient of the assessed characteristic by assuming that the more people recognize a person as displaying the assessed characteristic, the more prototypical the person is in regard to that characteristic. The two main methods for peer evaluation (see Kane & Lawler, 1978, for a review) share a similar logic. The first method, “peer rating,” consists of each group member rating his/her peers on a specific characteristic (e.g., contribution to the group), typically using a Likert scale for each characteristic. The final assessment score of an individual is calculated as the average of these evaluations (see, for example, Beehr, Ivanitskaya, Hansen, Erofeev, & Gudanowski, 2001; Klagge, 1995). In the “peer nominations” method (used in this study), each member of a group designates a specified number of group members as representing a high (above threshold) level of a particular characteristic (e.g., those they think have contributed the most to the group). The score of the assessment is calculated either as the sum of selections (i.e., the number of peers who selected the employee) or as the proportion of selections according to team size. Of these two measures, peer nomination is more widely used (Love, 1981) and more reliable and valid (Kane & Lawler, 1978; Shore et al.,

1992) than peer ratings. It seems that group members are better able to indicate a few group members who are high on an assessed dimension than to rate all the members of a group, perhaps because, given the limited choice, the prototypicality gradient is steeper, leaving less chance of error.

These two measures derive from the same logic. Since prototypicality is related to performance, leadership, and even emotional reactions (Eden & Leviatan, 1975; Hogg, 2001; Hogg & Terry, 2000; Hogg & van Knippenberg, 2003; Lord, Brown, Harvey, & Hall, 2001; Lord, Foti, & Devader, 1984); and since prototypicality has been shown to be stable over time (Epitropaki & Martin, 2004), especially in the same context (Brown & Lord, 2001), we argue that individuals perceived initially as prototypical will continue to be perceived thus, and will therefore perform better in the future. Thus, the logic of the peer-nominations measure is that individuals who are more frequently nominated or receive higher rating are considered as more prototypical in regard to the assessed characteristic, and are therefore likely to perform better.

This theoretical link proposed between peer nominations and future performance has been substantiated by Moreno’s “sociometry.” Moreno identified individuals selected by many of their peers, suggesting that they were higher in the “group hierarchy” than others, and can therefore be classified as better performers according to the peer-nominations measure (Moreno, 1934). Building on the established fact that level of assessment in the present is a valid predictor of an individual’s level of performance in the future (Arthur, Day, McNelly, & Edens, 2003; Gaugler, Rosenthal, Thornton, & Bentson, 1987), and on other research findings about the predictive validity of peer assessment measures (Kane & Lawler, 1978; Kraut, 1975; Love, 1981; Mayfield, 1970; Reilly & Chao, 1982; Roadman, 1964; Shore et al., 1992), we hypothesize that peer-nomination scores will predict future performance.

Hypothesis 1: Peer-nomination scores are positively related to future performance.

Utilizing insights gained from social information processing with analytic indicators derived from social network analysis of peer assessment data, we show that more complex measures of peer assessment should be considered. We believe that the peer-nominations measure may be inadequate for capturing the hierarchical nature of the prototypicality gradient of an assessed characteristic, because peer nomination discounts the prototypicality of the nominator. The measure is based on aggregation of nominations, thus assuming equality of nominees.

Nominations-Not-Returned as a Measure of Hierarchy

Individuals differ in the extent to which they display an assessed characteristic; hence, we argue that nominations-not-returned is an important indicator that taps into the prototypicality gradient. Louch (2000) found that individuals tend to return (or reciprocate) nominations, with about 75 percent of the selections in some networks exhibiting reciprocation. Thus, if *A* nominates *B*, in most cases *B* will also nominate *A*, reciprocity being a very powerful norm of human behavior (Cialdini, 2001).

However, according to theories such as social identity theory (Tajfel & Turner, 1979) and self-categorization theory (Turner, 1985; Turner et al., 1987), if *A* nominates *B* and *B* does not reciprocate, the nomination is meaningful in regard to the hierarchical position of *B* in the group, suggesting differentiation in the level of the assessed characteristic. Based on the social identity (Tajfel & Turner, 1979), self-categorization (Turner et al., 1987), and social comparison (Festinger, 1957) theories that group members conduct upward comparisons, we suggest that group members should *not* nominate other members less prototypical than themselves. Conversely, less prototypical individuals should select more prototypical members, thereby creating a hierarchical social structure via social categorization and upward comparisons.

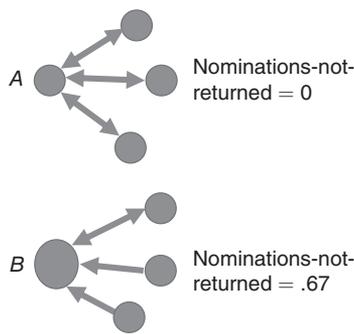
Reciprocal selections are indicators of equality (Hanneman & Riddle, 2005).

For example, in the case of peer assessment, returned nominations indicate that no member has a higher assessed characteristic than the nominating individual (e.g., both nominated each other as a contributor to the team). Nominations-not-returned, however, indicate levels of the assessment hierarchy where there are unequal assessments between two peers. These unequal assessments suggest a hierarchy of the assessed characteristic, and that social comparison is enhanced where there is a limited number of nominations (as in our sample), because the small number of members in the group makes it impossible to select all or even most of the peers.

We suggest that the nominations-not-returned measure captures those higher on the assessed characteristic, either because they are nominated by so many that they cannot reciprocate, due to limited number of nominations allowed, or because the nominated individuals select others as even higher on the assessed characteristic. Thus, in regard, for instance, to popularity, nominations-not-returned adds a hierarchy to the assessed characteristics based on dyadic comparisons while controlling for peer nominations.

Consider, for example, individuals *A* and *B*, both selected by three other peers as contributors to the team. Assume that each individual can only nominate three others. *A*'s and *B*'s peer-nominations score is exactly the same (3), since they both received three nominations. Then assume that *A* reciprocates by nominating everyone who has nominated him or her. In terms of the prototypicality gradient, *A* is at the same level as the individuals who nominated him. Thus, *A*'s nominations-not-returned score is 0. *B*, however, reciprocates only one nomination, and nominates two others who did not nominate him as contributing to the team. In this case, *B*'s nominations-not-returned score is 0.67, because, out of the three nominations, two

Based on the social identity, self-categorization, and social comparison theories that group members conduct upward comparisons, we suggest that group members should not nominate other members less prototypical than themselves.



Note: Two individuals with similar peer nominations (indegree = 3), but with different nominations-not-returned scores. Individual *B* is hierarchically higher than individual *A*. Circle size represents level of performance on the attribute.

FIGURE 1. Nominations-Not-Returned in Peer Assessment Network

were not returned. Accordingly, *B* is higher in the prototypicality gradient than one of the nominating individuals, and lower than the other two (see Figure 1).

We argue that, controlling for peer-nominations score and with a limited number of nominations, *B* is higher in the assessment hierarchy than *A*, and therefore has better potential to be a good performer in the future (see Figure 1). In sum, we posit that, controlling for incoming nominations, there will be a positive relationship between nominations-not-returned and future performance.

Hypothesis 2: Controlling for peer nominations, nominations-not-returned will add to the reliability of prediction of future performance. The more nominations individuals receive but do not return to the nominees, the higher their future performance.

Nominations by Nominees: The Hierarchy Concealed in Triads

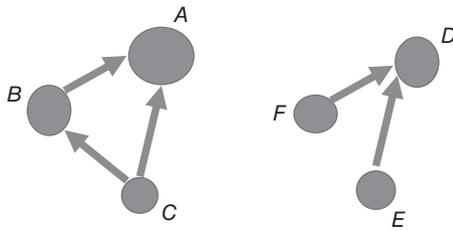
Sociologists theorize that triads—three actors and the linkages among them—are qualitatively different from dyads (Simmel, 1955). Moreover, insights derived from social network analysis suggest that examination only of dyadic measures (e.g., peer-nominations measure and/or nominations-not-returned) is

often inadequate for capturing network complexity and the effects on dependent variables (Burt, 1992; Robins, Pattison, Kalish, & Lusher, 2007).

The current measure of peer nominations determines one's position in a hierarchy according to the number of selections received from others, while the nominations-not-returned measure focuses on being selected without reciprocating the nominations. Both measures are dyadic, representing the hierarchical organization of the prototypicality gradient. However, hierarchy can also be captured by triadic measures. We suggest that nominations-by-nominees can provide additional information about an individual's hierarchical status in a network.

Nominations-by-nominees is defined thus: when peer *A* is the evaluated person, two peers, *B* and *C*, have nominated *A* on an assessed characteristic, and at least one of them has also nominated the other on that characteristic. Let us consider three instances. In the first case, *B* and *C* do not nominate each other on the assessed characteristic. In this case, *A* has a peer nomination score of 2 and a nominations-by-nominee score of 0, so *A* is clearly higher on the assessed characteristic hierarchy than *B* and *C*. In the second case, *A* is nominated by both *B* and *C*, but *C* also nominates *B*. In this case, *A*'s nominations-by-nominees score is 0.5. Thus, a hierarchy—*A* is higher than *B*, who is higher than *C*. The third case, in which *A* is nominated by *B* and *C*, who also nominate each other on the assessed characteristic, provides *A* with a nominations-by-nominees score of 1, since both of the two nominations between nominees exist (i.e., *A* was nominated by two individuals, both of whom display the assessed characteristic). In sum, *A* is higher in the prototypicality gradient, because he has been selected by prototypical others (see Figure 2).

Explanation of why a higher score in nominations-by-nominees in peer assessment indicates a higher level of prototypicality can be found in the literature about expert opinion (Brass & Burkhardt, 1993). Drawing on resource dependence, social comparison, and social information-processing theories,



Note: Nomination by nominees in peer nominations and performance. Individual A (evaluated in the left network) is nominated by individuals B and C as an expert, as does individual C; therefore, B is a nominee. Conversely, individual D (evaluated in the right network) is nominated by F and E (both of whom were not nominated by others). Based on the nominations-by-nominees measure, individual A is higher in the hierarchy than D although both received the same number of nominations (2). Circle size represents level of assessed performance on the attribute.

FIGURE 2. Nominations-by-Nominees in Peer Assessment Network

Brass and Burkhardt (1993) argued that an individual's centrality in the network of a group reinforces his/her reputational power. Individuals selected by other group members can be accepted as individuals with the reputation of experts and therefore better evaluators of a characteristic. For example, in assessments of cognitive ability, peers with the reputation of high cognitive ability should be better evaluators than those with lower cognitive ability. In other words, those who are higher on the assessed characteristic should be selected by more members of a group, but also by members selected by others as high on the characteristic.

Individuals perceived as prototypical should nominate others who are even more prototypical, according to social comparison theory (Festinger, 1957). A prototypicality hierarchy is thus created in which the more prototypical member is nominated by others who have themselves been nominated as prototypical.¹ The information obtained from this triadic comparison adds to information from the peer-nomination score, and from the proposed nominations-not-returned score that only concerns dyadic comparisons. We therefore hypothesize that nominations-by-nominees will improve prediction of future performance better than peer-nominations and nominations-not-returned:

Hypothesis 3: Nominations-by-nominees will improve prediction of future performance over other measures of assessment.

Expressive and Instrumental Assessments

Peer assessments often require employees to assess their peers according to several dimensions. In this study, we have assessed two psychological dimensions that frequently appear in studies on social information processing. Asch (1946) found that social information processing can be reduced to two primary dimensions: social and intellectual. Bales (1950) and Slater (1955) identified socioemotional orientation and task orientation as the two dimensions from which group members formed impressions of others. Leadership studies emphasize that subordinates evaluate managers on task and people orientation (Yukl, Gordon, & Taber, 2002). Fiske, Cuddy, and Glick (2007, p. 77) state that "people everywhere differentiate each other according to liking (warmth, trustworthiness) and respecting (competence, efficiency)."

In regard to social relationships, Ibarra (1992, 1993) indicates the universal dimensions of expressive and instrumental ties. Instrumental ties (in the context of work) involve exchange of job-related resources, including information, expertise, and material resources (Kram, 1988; Thomas, 1990). Expressive ties involve an exchange of friendly behavior and (emotional) support and are characterized by higher levels of trust and closeness than instrumental ties (Krackhardt, 1992).

In this study we concentrate on these two dimensions, referring to them as expressive versus instrumental dimensions (Ibarra, 1992, 1993). It is obvious that assessing others on the instrumental dimension is highly relevant to future performance, but why the expressive dimension is also relevant is less obvious. Casciaro and Lobo (2008) demonstrated that people prefer having instrumental ties with people they like. Social ability to connect with others helps individuals to tap into knowledge available in a group. Thus, individuals assessed as high on the expressive

dimension will ultimately perform better due to the resources acquired from cooperation with other group members. To assess how individuals perform and to validate the two peer-assessment dimensions, we focused on: engagement, performance under stress, and leadership. These are highly relevant to every work setting, and are distinct from one another, so that together they capture a wide spectrum of aspects of performance as related to peer assessment.

Social Network Analysis and Perceptions of Expressive and Instrumental Prototypical Behavior

Social network analysis is an analytical framework assisting data collection and analysis in order to capture organizational processes (Tichy, Tushman, & Fombrun, 1979). Social networks consist of actors, and of linkages that may be directed (i.e., have a source and a target). In the case of peer evaluations, the “actors” nominate others who display the assessed characteristic (linkages). This results in a network of nominations according to assessed characteristics (but *not* social relationships or ties between individuals). Such networks tap into prototypical representations of social reality (Freeman, 1992; Knoke & Kuklinski, 1982; Krackhardt, 1987). We selected a network-analysis framework because it offers indicators of hierarchy among nominees that comprise more information, and can capture the within-group prototype gradient.

We suggest that the new indicators—nominations-not-returned and nominations-by-nominees—can enhance our

knowledge of within-group social relationships, which can, in turn, add to our understanding about individual group members. Thus, future performance of an individual can be assessed not only by how many selections/nominations he or she

has received, but also by evaluating other indicators based on peer nominations that also reflect the hierarchy of nominations.

Method

Sample and Procedure

Participants were 249 Israeli male recruits, volunteers for a two-day selection process for an all-male military unit. All recruits were 18 years old and had completed high school. Selection took place out-of-doors and resembled selection processes in civil organizations, in which observers and raters evaluate candidates’ performance. Recruits were divided into 18 groups in which the commanders made sure that the recruits did not know each other. They were asked to perform a variety of tasks together with their new peers. As in other military tasks, the recruits were dependent on their peers in order to complete their assignments, and the group members were consistently together, making it possible for them to evaluate each other. Observers evaluated the recruits using code numbers printed on their shirts, which were visible throughout the procedure. The selection rate of recruits accepted to the unit was about 53 percent, and eventually 132 of these recruits were accepted to the unit.

At the end of the selection procedure, peer assessment was conducted in every group, focusing on evaluating friendly behaviors and contribution to the group, two aspects that are relevant indicators for elite unit performance and can be readily assessed after two days of shared activity. Each member of each group had to identify three other members who, in his opinion, contributed most to the group, and three that demonstrated more friendly behaviors during the selection than other group members.

The soldiers were placed in a circle to allow all group members to see the numbers printed on the shirts (they did not know each other by name). This made it possible for soldiers to see the numbers of the other group members, but because they continuously scanned the entire circle, it was difficult to note who selected

Future performance of an individual can be assessed not only by how many selections/nominations he or she has received, but also by evaluating other indicators based on peer nominations that also reflect the hierarchy of nominations.

whom, or for them to look at each other's assessments. A member of the psychology department distributed the questionnaires, explaining that the recruits should write the numbers of those they nominated, based on their reactions to these members' behavior during the two previous days. For each group, recruits' responses were entered into two matrices, so that a "1" in cell- i,j indicates that individual i nominated individual j as representing the assessed characteristic. One such matrix represented assessment of friendly behavior (i.e., our "friendly behavior network"). The other represented assessment of the contribution to the team (i.e., our "contribution network")

Each unit commander also evaluated his recruits, on a Likert scale, after interviewing each of them at the end of the selection process.

To validate the peer-assessment tool, we conducted a performance assessment of these soldiers about six months later. Performance was assessed by the commanders who had been working with them and training them. It is important to note that the commanders acted as guides during the basic training period. That is, they were consistently with the soldiers, giving them assignments of increasing difficulty, observing their performance, and providing feedback on their evaluated performance. The commanders also collected multiple measures reflecting each soldier's behavior, written tests about weapons and tools, objective measures of performance such as number of hits in target practice, or time taken for an obstacle course. Based on their evaluations, the commanders assessed the soldiers on a Likert scale, on three measures that are central to a soldier's role: performance under stress, engagement, and leadership.

Measures: Peer Assessment

The *peer-nominations score* represents the number of people who select an individual as demonstrating friendly behaviors (in the friendly behavior network) or as a contributor to the team (in the contribution network). Since the groups differed slightly in size, we normalized indegree centrality, calculating it

as the proportion of others on the team who nominated the respondent as demonstrating friendly behavior or as a contributor (Wasserman & Faust, 1994). The index ranged from 0 (the respondent was not selected) to 1 (the respondent was selected by everyone on the team). This represents the traditional organizational measure in peer-assessment tools, and is similar to the measure called *indegree centrality* in social network analysis.

Nominations-not-returned was calculated as the proportion of unreciprocated selections (in both networks) from other group members. Nominations-not-returned scores ranged from 0 (all those who selected an individual were reciprocally selected) to 1 (none of those who selected an individual were reciprocally selected by that individual). This is similar to the nonreciprocity measure in social network analysis.

Nominations-by-nominees was calculated as the number of selections (in both networks) of an individual divided by the potential number of selections among others and presented in percentages. Thus, our measure of nominations-by-nominees is $d = 2l/n(n-1)$, where n is the number of peer nominations and l reflects the number of selections of an individual by others in the friendly behavior and/or contribution network. Since we define n according to peer nominations (the number of selections received), our nominations-by-nominees measure is similar to the more traditional measure of (indegree-based) egocentric network density. Nominations-by-nominees ranges from 0 (none of the group members who selected the individual selected each other) to 1 (all of the group members also selected each other). All three peer-assessment measures were calculated using the UCINET program (Borgatti, Everett, & Freeman, 2002).

A *supervisor's assessment* (control) tested whether the information collected from peers

It is important to note that the commanders acted as guides during the basic training period. That is, they were consistently with the soldiers, giving them assignments of increasing difficulty, observing their performance, and providing feedback on their evaluated performance.

during the selection process added to the prediction of future performance beyond the traditional, subjective top-down assessment of a single supervisor. Our control derived from the assessment of a commander of the unit who interviewed each soldier during the selection process in order to evaluate the soldier’s suitability for the unit. It is important to note that the supervisor’s assessment was conducted at the end of the selection process in parallel to the peer assessment (and six months prior to the assessment of performance).

Performance after six months was assessed by two to four direct commanders of each soldier for about half a year (after the selection) and were familiar with the soldiers’ performance (they had also monitored, rated, graded, and documented the performance of each soldier over a six-month period). In order to rate performance level, the commanders used a Likert scale, ranging from 1 (very low) to 5 (very high). Evaluations were of three key performance dimensions:

The ability to command is an important component of a soldier’s performance. Many potential leaders are required for the success of a military unit, for which reason all soldiers are given formal leadership training, and leadership performance is evaluated by their commanders

1. *Engagement*: defined as “high levels of activity, initiative, and responsibility” (Dvir, Eden, Avolio, & Shamir, 2002, p. 737). Soldiers need to be committed and highly motivated toward developing fighting skills and abilities. Investing effort in achieving this goal in the first six months of military training is a key element of success and is very evident to their commanders. The commanders collected information about engagement behaviors for each soldier. For example, engaged soldiers are more likely to volunteer for tasks in their free time that demand extra effort (e.g., carrying/being in charge of special equipment such as communication equipment, heavy machine guns, etc.). Performance of such extracurricular tasks is documented. Summarizing such indicators of engagement informed command-

ers’ ratings of each soldier on the five-point Likert scale.

2. *Performance under stress*: stressful, threatening, or demanding situations can be conducive to a decrease in performance (Driskell & Salas, 1991, 1996). The soldier’s role in a combat unit is difficult, comprising mental and physical challenges in extreme conditions “that may result in extensive . . . physical, psychological, or material consequences. . . .” (Hannah, Uhl-Bien, Avolio, & Cavarretta, 2009, p. 898). Soldiers’ ability to perform under such stress and in life-threatening conditions is essential, which is why military training is intentionally stressful and demanding (Luria & Torjman, 2009), enabling commanders to evaluate performance in different stressful circumstances. For example, shooting ability and accuracy is measured over six months by counting the number of direct hits in shooting practices, under varying levels of stress. The level of accuracy of shooting would not decrease much in stressful exercises if the soldiers performed well under stress.
3. *Leadership*: a mechanism of interpersonal influence over subordinates (e.g., Yukl, 2002). The ability to command is an important component of a soldier’s performance. Many potential leaders are required for the success of a military unit, for which reason all soldiers are given formal leadership training, and leadership performance is evaluated by their commanders. For example, each soldier is assigned as “soldier on duty” for several days of the basic training period. This role includes responsibility for maintaining a department schedule defined by the commanders. It also includes managing departmental logistic duties and tasks, and the division of jobs among the soldiers (guard duty, kitchen duties, etc.). The commanders monitor how well these tasks are accomplished by the department soldiers.

In our study, the unit commanders were unaware of the research teams’ hypotheses, of candidates’ scores of half a year earlier, and

of other commanders' ratings of the soldiers' performance. In order to assess inter-rater agreement among these observers, we used a special Intra-class Correlation known as ICC [1,k] (see Shrout & Fleiss, 1979, for a similar procedure). This form of ICC is adjusted for varying numbers of raters for different participants. The average ICC (1,k) assessment was .829 for engagement, .873 for leadership, and .819 for performance under stress, indicating good agreement among raters. The mean of the ratings of all observers was used to calculate interdependent performance scores.

Results

Means, standard deviations, and correlations among variables are presented in Table I. Correlations among peer assessment measures (peer nominations, nominations-by-nominees, nominations-not-returned) of the two assessed networks (friendly behavior and contribution) are not high (average = .36), suggesting that these are two discrete domains of assessment. Furthermore, correlations among these measures within each network are even smaller (average = .225 for friendly behavior and .275 for contribution), demonstrating that the three measures also measure properties specific to each network.

We found that supervisors' assessment of soldiers during the selection (interview score) was related to the peer-nomination score on both dimensions, as well as to the nominations-by-nominees score on the friendly behavior network, but we found no significant relationship between supervisors' assessment of the soldiers during selection and the soldiers' performance half a year later. We did, however, find some significant relationships among peer-assessment measures in the selection and the soldiers' performance half a year later.

The significant correlations of network measures with performance suggest that these measures may add to prediction of future performance. However, in order to determine if any of these measures predicts performance better than others, especially in regard to peer nominations, we tested

our hypotheses in a combined, hierarchical linear model (HLM), which included all of the predictors. HLM is suitable for testing relations when individual data are nested within groups (Bryk & Raudenbush, 1992). With three dependent variables (engagement, leadership, and performance under stress) and two assessment dimensions (contribution and friendly behavior), we ran six regressions, each regression predicting a dependent variable from one network.

The results of the HLM analysis, presented in Table II for the friendly behavior network and Table III for the contribution network, support our hypotheses, but only for one specific network and dimension. Hypothesis 1 suggested that the peer-nominations score would be related to future performance. As shown in Table III, there was only weak support for this hypothesis, since the score was only related to engagement, and only in the contribution-to-the-group assessment. Neither of the other dimensions (performance under stress and leadership) was significantly related with peer nominations in either of the assessed networks.

Hypothesis 2 was that nominations-not-returned would predict performance. As shown in Table II, nominations-not-returned did not correlate significantly with any of the three dimensions of performance (engagement, leadership, and performance under stress) in the friendly behavior network, after controlling for peer-nominations score and all other variables. This was the only measure significantly related to centrality in the friendly-behavior network. On the other hand, it was not significantly related to performance in contribution-to-the-group assessment (see Table III).

Hypothesis 3 suggested that the nominations-by-nominees score would add to the prediction of performance. As shown in Table III, nominations-by-nominees significantly predicted engagement and performance under stress after controlling for supervisors' evaluation, peer-nominations, and nominations-not-returned scores in the contribution to the group assessment network, but did not predict performance in the friendly behavior network (see Table II). In sum, there was support

T A B L E I Means, Standard Deviations, and Intercorrelations Among Variables

	1	2	3	4	5	6	7	8	9	10
1. Peer nominations—FB										
2. Peer nominations—Contribution	.665***									
3. Nominations-by-nominees—FB	.243***	.108								
4. Nominations-by-nominees—Contribution	.422***	.382***	.225***							
5. Nominations-not-returned—FB	.381***	.297***	.053	.118						
6. Nominations-not-returned—Contribution	.180**	.386***	.063	.059	.191**					
7. Interview	.234***	.202**	.194**	.059	.092	.105				
8. Engagement	.109	.227**	-.069	.254**	.274**	.005	-.060			
9. Stress Performance	.044	.147	-.078	.195*	.242**	-.011	.032	.668***		
10. Leadership	.130	.176*	-.107	.120	.226**	.095	.120	.531***	.676***	
<i>N</i>	249	249	249	249	249	249	249	132	132	132
Mean	.214	.213	17.80	12.31	.447	.515	69.88	3.879	3.923	3.498
<i>SD</i>	.154	.199	22.79	18.36	.342	.378	14.79	.918	.807	.992

Notes: FB = friendship behavior network; Contribution = contribution to the group network.
p* < .05, *p* < .01, ****p* < .001.

TABLE II Hierarchical Linear Modeling Analysis Predicting Performance From Indicators of the Friendship Behavior Network

Variable	Leadership			Performance Under Stress			Engagement		
	Coefficient	SE	T Ratio	Coefficient	SE	T Ratio	Coefficient	SE	T Ratio
Level 2									
Intercept	3.220***	.194	16.54	3.72***	.164	22.636	3.562***	.184	19.271
Level 1									
Interview	.0003	.008	.039	-.004	.006	-.720	-.012	.006	-1.739
Peer nominations	.345	.594	.581	-.084	.479	-.17	.198	.525	.378
Nominations-by-nominees	-.004	.003	-1.372	-.002	.002	-.815	-.001	.003	-.637
Nominations-not-returned	.639*	.295	2.164	.601*	.23	2.53	.717**	.26	2.758

Notes: Level 1: n = 132, Level 2: n = 18, Interview = supervisor ratings during selection process.

*p < .05, **p < .01, ***p < .001.

TABLE III Hierarchical Linear Modeling Analysis Predicting Performance From Indicators of Contributions to the Group Network

Variable	Leadership			Performance Under Stress			Engagement		
	Coefficient	SE	T Ratio	Coefficient	SE	T Ratio	Coefficient	SE	T Ratio
Level 2									
Intercept	3.173***	.193	16.433	3.762***	.166	22.66	3.634***	.179	20.226
Level 1									
Interview	-.001	.008	-.09	-.005	-.006	-.867	-.128	.006	-1.917
Peer nominations	.611	.486	1.257	.504	.372	1.357	.894*	.407	2.196
Nominations-by-nominees	.004	.005	.995	.009*	.003	2.331	.009*	.004	2.249
Nominations-not-returned	.142	.275	.518	-.202	.219	-.926	-.216	.239	-.906

Notes: Level 1: n = 132, Level 2: n = 18, Interview = supervisor ratings during selection process.
*p < .05, ***p < .001.

for Hypothesis 3 in the contribution network but not in the friendly behavior network.

Discussion

The aim of this study was to test the contribution of additional network analytic measures examining hierarchy to the predictive validity of peer assessment. Longitudinally, we found that nominations-by-nominees was the best predictor of performance in the peer-assessment network of instrumental contribution to the group on the dimensions of engagement and performance under stress, but not in leadership. On the other hand, nominations-not-returned was the only measure to predict performance in the peer-assessment network of friendly behavior (for all three dimensions). The peer-nominations score, the only measure traditionally used in peer assessment, was found to be a valid predictor of performance (together with nominations-by-nominees), but only in assessment of contribution to the group and only for engagement. It was also found to relate to supervisors' assessment during the selection process (see Table IV). Thus, it seems that nominations-by-nominees and nominations-not-returned do in fact improve prediction of performance, and were found to be better predictors of performance than traditional assessment measures such as peer nominations and supervisors' assessment. It also appears that supervisors react to the number of peer nominations but not to other indicators that have higher predictive validity.

We found that peer nominations predicted performance only in regard to contribution to group assessment together with nominations-by-nominees. It is possible that this is due to the complexity of assessing contribution to the group (versus friendly behavior), and that for such more complex assessments multiple measures are needed. We think that peer nominations only predicted engagement (and not performance under stress or leadership) because of the theoretical proximity of contribution to the group and engagement, and that highly engaged individuals (i.e., those demonstrating high levels of activity and responsible behavior) are prototypical

contributors to the group. Measuring three different dimensions of performance that are relevant to work settings and distinct from each other (i.e., engagement, performance under stress, and leadership) provided the first empirical results about what exactly peer-assessment measures can predict. Many human resource practices aim to achieve a higher level of these dimensions of performance, and this study's contribution is the provision of empirical evidence concerning which peer-assessment measures are related to each of these dimensions.

In regard to the more complex measures and future performance, we found that for each assessed network (friendly behavior, contribution to the group), only one of these measures is valid in prediction of performance (i.e., validity of peer-assessment measures depends on the content of the dimension assessed). Although we hypothesized that nominations-not-returned and nominations-by-nominees would be valid predictors for instrumental and expressive assessment networks alike, it appears that a different, content-related logic operates for each network.

The distinction between the two assessments used in this study (friendly behavior and contribution) is aligned with the distinction between instrumental and expressive network ties in the literature (see Fombrun, 1982; Ibarra, 1992, 1993), a recurring categorization of how individuals perceive others (Asch, 1946; Bales, 1950; Fiske et al., 2007; Slater, 1955). Apparently, this fundamental categorization was not considered in analyzing human resource practices. It is possible that practices varying from selection of potential employees to their assessment and development may be different when separating instrumental and expressive content. We found no discussion of the effects of assessment content on network measures, and

It seems that nominations-by-nominees and nominations-not-returned do in fact improve prediction of performance, and were found to be better predictors of performance than traditional assessment measures such as peer nominations and supervisors' assessment.

T A B L E I V		Summary of Results: Network Indicators for Future Performance—Friendship Behaviors and Contribution to the Group Assessment Networks	
	Engagement	Performance Under Stress	Leadership
<i>Friendship Behavior Network</i>	Nominations-not-returned	Nominations-not-returned	Nominations-not-returned
<i>Contribution to the Group Network</i>	Peer nominations Nominations-by-nominees	Nominations-by-nominees	

therefore can only rely on other bodies of literature that differentiate among these aspects.

It is possible that the difference between predictors of performance in instrumental and emotional assessment networks derives from the prototype being assessed. We believe that the difficulty of assessing others on a prototype and the complexity of evaluation may influence categorization. Rosch and Mervis (1975) found that not all prototypes are so easily memorized or so informative for guidance. Others (see Kunda, 2002) suggest that prototypical thinking relates to environmental factors such as social complexity. Harrison, Price, Gavin, and Florey (2002) suggested that, in evaluating others, social information processing is affected by two sources of diversity between individuals. Surface diversity reflects visible characteristics that are easily observed and measured (such as age, sex, or race). Deep-level diversity is less apparent and includes personality traits and values, as well as attitudes, preferences, and beliefs (see also Harrison, Price, & Bell, 1998). We suggest that the instrumental contribution to group-dimension assessment involves more complex, deep-level prototypes; hence, there is greater need for evaluation of experts/prototypical members within a group (i.e., nominees). Conversely, the expressive prototype is less complex and would not require consulting with others in regard to assessment. The dyadic comparison between pairs (nominations-not-returned) captures those of the group who behave more expressively than their peers. The paired comparisons among members of a group clearly identify those that behaved in a friendly manner (i.e., the highly expressive prototypical members).

The dyadic measure of nominations-not-returned points to group members assessed as those that behaved in a friendly manner in the selection and therefore are likely to receive support from others who view them as friends. Over time, they may even receive more support from others because they are more likely to build social ties with others. This suggests that prototypicality—high scores on the different indicators derived from the assessment network—is not only related to performance due to group

hierarchy but also because it becomes an actual resource. In other words, individuals behaving prototypically may ultimately acquire social capital (i.e., the “ability to secure benefits through membership in networks and other social structures”; Portes, 1998, p. 8). It is possible that group members displaying prototypical behavior will achieve social support that is less available to less prototypical members, and will perform better due to their accumulated social capital.² Thus, social capital may develop after selection of prototypical members and mediate the relationship between prototypicality and future performance.

We also found that while nominations-not-returned in the friendly behavior network predicted all three dimensions of performance (engagement, performance under stress, and leadership), nominations-by-nominees predicted two out of the three measures in contribution to group assessment but did not predict leadership performance. The fact that a dyadic (as opposed to triadic) hierarchy measure reflects leadership is entirely consistent with research findings that the hierarchical structure of emergent leadership in the military is primarily dyadic (Kalish & Luria, 2013).

In our review of peer assessment studies, we found none that integrate the social cognition that occurs during peer assessment with assessment outcomes and measures. Peer assessment can be explained by social identity and self-categorization theories (Tajfel & Turner, 1979; Turner et al., 1987), providing a bridge between applied practices of human resource and social theories. We believe that peer appraisal is an element of real life that can be studied and tested in order to improve social theories. Great advantages can be derived from deeper theoretical understanding of peer-assessment procedures. As seen in this study, theoretical understanding can lead to new indicators for improving predictive validity.

Our results demonstrate that to enhance predictive validity (at least in this sample), the measure traditionally used in peer assessment (peer nominations) should be supplemented with the more complex measures of nominations-not-returned and

nominations-by-nominees. Logically, these more complex measures convey significant information about the hierarchy of the prototypicality gradient, and hence of the peer-assessment network. We also found that supervisors’ assessments are mostly correlated with peer-nomination scores (calculated on the basis of the peer-assessment data). Our results indicate that the predictive validity of peer assessment can only be improved by using a different calculation of the data collected during the traditional peer-assessment procedures, and that these improvements will add to the prediction value of supervisory assessments.

Such assessments are also used in organizations for employee development (Farh, Cannella, & Bedeian, 1991). Developmental evaluations are based on feedback (Brett & Atwater, 2001). Peer-assessment scores (peer nominations) were assigned in order to identify areas of performance that needed improvement. It is possible that returned nominations-by-nominees provide additional information that enhances an individual’s understanding of his or her performance in the group. Our results show improved predictive validity of the proposed measures, which is of the utmost importance for such feedback-based practices. More accurate ratings in human-resource feedback processes lead to increased trust and are even linked to organizational effectiveness outcomes (Bettenhausen & Fedor, 1997; Drenth, 1998; Farh et al., 1991). According to Drenth (1998), lack of accuracy in rating scores damages the fairness perception of employees.

Our results demonstrate that not all dimensions of assessment relate similarly to future performance, so that different measures may be valid for different dimensions of assessment. By means of three measures of assessment (peer nominations, nominations-by-nominees, and nominations-not-returned), practitioners can test which measures are most relevant to each assessed

Our results demonstrate that not all dimensions of assessment relate similarly to future performance, so that different measures may be valid for different dimensions of assessment.

dimension. Hence, due to a simple validation process, practitioners can use the most appropriate measures and improve the accuracy of their assessment scores. In our study, we recommended that the military human resource practitioner should use the nominations-not-returned measure to calculate the friendly-behaviors assessment, and the nominations-by-nominees measure to calculate the contribution assessment.

We note the limitation of using only top-down assessment of the dependent variable, and suggest that future research should employ objective measures as the dependent variables, as well as, where possible, multiple sources of evaluation.

Our results may also be applied to manager training and development programs, a key element of which is creating a mechanism for obtaining information that managers do not otherwise receive about aspects of their behavior (London & Smither, 1995; Waldman, Atwater, & Antonioni, 1998). Our results demonstrate that manager assessments are correlated with traditional peer-nomination scores, and not with the more complex measures (nominations-not-returned and nominations-by-nominees). Thus, managers seem to only note dimensions of performance captured by the peer nominations, rather than noticing more valid information revealed by the more complex measures. This is aligned with studies by Janicik and Larrick (2005) who demonstrated that accurate perception of networks, especially of more complex network structures, is not intuitive but can be learned. We suggest that managers should be

trained to analyze data about employees based on complex measures of peer assessment for this purpose. Such information was shown in this study to be related to future performance, and therefore should improve managerial decisions regarding subordinates.

Finally, it is important to note that our peer-assessment process was based on a fixed number of nominations. We suggest that the complex measures tested in this study should also apply to assessment processes based on an unlimited number of nominations. Since our measures tap into the hierarchical structure

of the prototypicality gradient, they should be relatively independent of respondents' freedom to nominate as many others as they wish.³ However, based on the literature of prototypicality discussed in this study, we do find it advantageous to fix the number of nominations, since nominating a small number would allow every participant to point to those who fit the prototype exceptionally well, thereby making assessment easier and, hence, reduce errors of assessment.

Limitations

One limitation of this study is the unique sample of 18-year-old male soldiers. As a selected group from a specific culture with relatively similar backgrounds, this is a homogeneous sample that makes it possible to control for potential intervening variables, but also limits the generalizability of the findings. Previous studies (Dvir et al., 2002; Shamir, Zakay, Breinin, & Popper, 1998) discussed the similarities between civilian and military contexts, so we do not anticipate that the results will be dependent on the sample. Future research should extend the examination of these network measures to nonmilitary settings.

A second limitation is related to conducting peer assessment within a real-life selection process. The process may have biased individuals' responses to surveys due to social desirability. Despite this limitation, meta-analytic findings suggest that while social desirability does play a role in personnel selection, it is "not a pervasive problem as had been anticipated by IO psychologists" (Ones, Viswesvaran, & Reiss, 1996, p. 660). Future research should test the contribution of these network measures to ongoing organizational assessment, both in terms of their validity and the value of the additional information to feedback and improvement based on the measurement.

A third limitation is that cross-sectional design precluded the possibility of determining causality. Although the design was longitudinal, used terms such as predictive validity, and hypothesized that peer-assessment measures can predict performance, it is important to note that, because this is far from being a

controlled experiment, we cannot conclude that high scores in peer assessment are conducive to better performance.

Our criteria for examining the predictive validity of our measures were based on commanders' assessment of the soldiers' performance during their basic training. This measure has advantages over regular supervisory assessment in civil organizations because the commanders consistently monitor soldiers' performance during basic training and use multiple objective tests in their evaluations, as explained in the Method section. Furthermore, each soldier was assessed by several commanders, with high reliability between their evaluations. There are, however, criticisms in the literature of the classic supervisory-based assessment, which has been shown to be less valid than peer assessment (Kane & Lawler, 1978). Although the commanders' assessment has many advantages over the classic supervisors' assessment measurement in civil organizations, we note the limitation of using only top-down assessment of the dependent variable, and suggest that future research should employ objective measures as the dependent variables, as well as, where possible, multiple sources of evaluation.

Future Research Agenda

Our results show that performance is predicted by structural characteristics of an individual's position in the assessment network. Future research should investigate personal characteristics such as traits and predispositions of individuals who are nominated but do not return nominations, and of those who are nominated by nominees (Kalish & Robins, 2006; Klein, Lim, Saltz, & Mayer, 2004). We suggest that future studies should also compare multidimensional feedbacks, including nominations-by-nominees, nominations-not-returned, and peer nominations, with the current single-dimension feedback based only on peer nominations. Such studies should also test validity and accuracy of unlimited nominations for peer-assessment processes. Adding an independent measure of prototypicality to peer-nomination

measurements may provide further support for our theoretical explanation.

Our results indicate that different networks follow different paths of logic (Rank, Robins, & Pattison, 2010). For some, such as the friendly behavior assessment, future performance is related to nominations-not-returned; for others, such as contribution assessment, future performance is related to nominations-by-nominees. Assessments and network studies should therefore consider the assessed dimension and the "structural logic" applicable to each network individually.

Acknowledgments

We wish to thank the anonymous reviewers for their comments and feedback during the review process, which improved our manuscript considerably. We also wish to thank Major Miriam Weinstein and Allon Kahanaor for their support in executing this research.

Notes

1. We note that there could be multiple hierarchies of prototypicality, in which different individuals choose prototypical members based on different norms. The deviating norms should add noise to the peer-assessment measure regardless of what calculations are used. Positive correlations between the peer-assessment measure and future performance in this and previous studies indicate that the phenomenon is infrequent.
2. An alternative explanation to the relationship between our measures and performance is that the measures tap social capital rather than prototypicality. However, in this setting, it is unlikely that any participant had higher social capital during the selection because of the short period in which these group members worked together, and their isolation from other social networks. Thus, in this specific "field laboratory," peer ratings would not have been influenced by social capital.
3. Our measures are logically similar to those that examine hierarchy and power in networks with unlimited nominations. The nominations-not-returned measure is similar to Krackhardt's (1994) measures of least-upper-boundedness, while the nominations-by-nominees measure is similar to Bonacich's (1987) power.

GIL LURIA is a senior lecturer in the faculty of social welfare and health sciences in the Department of Human Services at the University of Haifa. He conducts research on organizational climate, leadership, stress, and organizational interventions. He received his PhD from the Faculty of Management at the Technion—Israel Institute of Technology. His studies have been published in scientific and professional journals, and adopted by practitioners in a number of organizations. His recent work focuses on social network analysis, employee performance, and behavior in organizations, with an emphasis on safety and service quality.

YUVAL KALISH is a lecturer in the Leon Recanati Graduate School of Business Administration at Tel Aviv University. He received his PhD from the Department of Psychology at the University of Melbourne. His research, which has been published in scientific and professional journals, focuses on social networks and individual attributes, mathematical models of social networks, leadership, and leadership emergence. His recent work focuses on negative and ambivalent relationships and on the interplay among social networks, context, and individual psychological attributes.

References

- Allen, T. J., & Cohen, S. I. (1969). Information flow in research and development laboratories. *Administrative Science Quarterly*, 14, 12–19.
- Ancona, D., & Caldwell, D. F. (1992). Bridging the boundary: External activity and performance in organizational teams. *Administrative Science Quarterly*, 37, 634–665.
- Arthur, W., Jr., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, 56, 125–154.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal Social Psychology*, 41, 258–290.
- Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups*. Reading, MA: Addison-Wesley.
- Beehr, T. A., Ivanitskaya, L., Hansen, C. P., Erofeev, D., & Gudanowski, D. M. (2001). Evaluation of 360 degree feedback ratings: Relationships with each other and with performance and selection predictors. *Journal of Organizational Behavior*, 22, 775–788.
- Bettenhausen, K. L., & Fedor, D. B. (1997). Peer and upward appraisals: A comparison of their benefits and problems. *Group Organization Management*, 22, 236–263.
- Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92, 1170–1182.
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCINET for Windows: Software for social network analysis*. Lexington, KY: Analytic Technologies.
- Borman, W. C. (1974). The rating of individuals in organizations: An alternative approach. *Organizational Behavior and Human Performance*, 12, 105–124.
- Brass, D. J., & Burkhardt, M. E. (1993). Potential power and power use: An investigation of structure and behavior. *Academy of Management Journal*, 36, 441–470.
- Brett, J. F., & Atwater, L. E. (2001). 360° feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology*, 85, 930–942.
- Brewer, M. B., & Harasty, A. S. (1996). Seeing groups as entities: The role of perceiver motivation. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition*, Vol. 3: The interpersonal context (pp. 347–370). New York, NY: Guilford.
- Brown, D. J., & Lord, R. G. (2001). Leadership and perceiver cognition: Moving beyond first order constructs. In M. London (Ed.), *How people evaluate others in organizations* (pp. 181–202). Mahwah, NJ: Lawrence Erlbaum.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, 3, 14–25.
- Casciaro, T., & Lobo, M. S. (2008). When competence is irrelevant: The role of interpersonal affect in task-related ties. *Administrative Science Quarterly*, 53, 655–684.
- Cialdini, R. B. (2001). *Influence: Science and practice* (4th ed.). Boston, MA: Allyn & Bacon.
- Colarelli, S. M., & Boos, A. L. (1992). Sociometric and ability-based assignment to work groups: Some implications for personnel selection. *Journal of Organizational Behavior*, 13, 187–196.
- Drenth, P. J. D. (1998). Personnel appraisal. In P. J. D. Drenth, H. Thierry, & C. J. de Wolff (Eds.), *Handbook of organizational psychology: Personnel psychology* (pp. 59–88). East Sussex, UK: Psychology Press.
- Driskell, J. E., & Salas, E. (1991). Group decision making under stress. *Journal of Applied Psychology*, 76, 473–478.
- Driskell, J. E., & Salas, E. (1996). *Stress and human performance*. Mahwah, NJ: Lawrence Erlbaum.
- Dvir, T., Eden, D., Avolio, B. J., & Shamir, B. (2002). Impact of transformational leadership on follower development and performance: A field experiment. *Academy of Management Journal*, 45, 735–744.
- Eden, D., & Leviatan, U. (1975). Implicit leadership theory as a determinant of factor structure underlying supervisory behavior scales. *Journal of Applied Psychology*, 60, 736–741.
- Epitropaki, O., & Martin, R. (2004). Implicit leadership theories in applied settings: Factor structure, generalizability, and stability over time. *Journal of Applied Psychology*, 89, 293–310.
- Farh, J. L., Cannella, A. A., Jr., & Bedeian, A. G. (1991). Peer rating: The impact of purpose on rating quality and user acceptance. *Group & Organization Management*, 16, 367–386.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11, 77–83.
- Fombrun, C. (1982). Environmental trends create new pressures on human resources. *Journal of Business Strategy*, 3(1), 61–69.
- Freeman, L. C. (1992). Filling in the blanks: A theory of cognitive categories and the structure of social affiliation. *Social Psychology Quarterly*, 55, 118–127.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72, 493–511.
- Geoffrey, B. (1994, November). Leader of the pack. *Sales and Marketing Management*, pp. 32–38.
- Hamilton, D. L., Sherman, S. J., & Lickel, B. (1998). Perceiving social groups: The importance of the entitativity continuum. In C. Sedikides, J. Schopler, & C. A. Insko (Eds.), *Intergroup cognition and intergroup behavior* (pp. 47–74). Mahwah, NJ: Lawrence Erlbaum.
- Hannah, S. T., Uhl-Bien, M., Avolio, B. J., & Cavarretta, F. L. (2009). A framework for examining leadership in extreme contexts. *Leadership Quarterly*, 20, 897–919.
- Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California, Riverside.
- Harrison, D. A., Price, K. H., & Bell, M. P. (1998). Beyond relational demography: Time and the effects of surface- and deep-level diversity on work group cohesion. *Academy of Management Journal*, 41, 96–107.
- Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning. *Academy of Management Journal*, 45, 1029–1045.
- Haslam, S. A., Oakes, P. J., McGarty, C., Turner, J. C., & Onorato, S. (1995). Contextual changes in the prototypicality of extreme and moderate outgroup members. *European Journal of Social Psychology*, 25, 509–530.
- Hogg, M. A. (1992). *The social psychology of group cohesiveness: From attraction to social identity*. London, UK: Harvester Wheatsheaf.
- Hogg, M. A. (1993). Group cohesiveness: A critical review and some new directions. *European Review of Social Psychology*, 4, 85–111.
- Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review*, 5, 184–200.
- Hogg, M. A., Cooper-Shaw, L., & Holzworth, D. W. (1993). Group prototypicality and depersonalized attraction in small interactive groups. *Personality and Social Psychology Bulletin*, 19, 452–465.
- Hogg, M. A., & Hains, S. C. (1998). Friendship and group identification: A new look at the role of cohesiveness in groupthink. *European Journal of Social Psychology*, 28, 323–341.
- Hogg, M. A., & Terry, D. J. (2000). Social identity and self-categorization processes in organizational contexts. *Academy of Management Review*, 25, 121–140.

- Hogg, M. A., & Terry, D. J. (2001). *Social identity processes in organizational contexts*. Philadelphia, PA: Psychology Press.
- Hogg, M. A., & van Knippenberg, D. (2003). Social identity and leadership processes in groups. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 1–52). San Diego, CA: Academic Press.
- Hornsey, M. J., & Hogg, M. A. (1999). Subgroup differentiation as a response to an overly-inclusive group: A test of optimal distinctiveness theory. *European Journal of Social Psychology*, 29, 543–550.
- Hornsey, M. J., & Hogg, M. A. (2000). Subgroup relations: A comparison of the mutual intergroup differentiation and common ingroup identity models of prejudice reduction. *Personality and Social Psychology Bulletin*, 26, 242–256.
- Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly*, 37, 422–447.
- Ibarra, H. (1993). Personal networks of women and minorities in management: A conceptual framework. *Academy of Management Review*, 18, 56–87.
- Janicik, G. A., & Larrick, R. P. (2005). Social network schemas and learning of incomplete networks. *Journal of Personality and Social Psychology*, 88, 348–364.
- Kalish, Y., & Luria, G. (2013). Brain, brawn, or optimism? Structure and correlates of emergent military leadership. In D. Lusher, J. Koskinen, & G. Robins (Eds.), *Exponential random graph models for social networks: Theories, methods and applications* (pp. 226–235). Cambridge, UK: Cambridge University Press.
- Kalish, Y., & Robins, G. (2006). Psychological predispositions and network structure: The relationship between individual predispositions, structural holes and network closure. *Social Networks*, 28, 56–84.
- Kane, J. S., & Lawler, E. E., III. (1978). Methods of peer assessment. *Psychological Bulletin*, 85, 553–586.
- Kane, J. S., & Lawler, E. E., III. (1979). Performance appraisal effectiveness: Its assessment, and determinants. In B. M. Staw (Ed.), *Research in organizational behavior* (pp. 425–478). Greenwich, CT: JAI Press.
- Katzenbach, J. R., & Smith, D. K. (1993). The discipline of teams. *Harvard Business Review*, 71(2), 111–120.
- Klagge, J. (1995). 360-degree sociometric feedback for individual and organizational change. *Public Administration Quarterly*, 19, 352–367.
- Klein, K. J., Lim, B. C., Saltz, J. L., & Mayer, D. M. (2004). How do they get there? An examination of the antecedents of centrality in team networks. *Academy of Management Journal*, 47, 952–963.
- Knoke, D. K., & Kuklinski, J. H. (1982). *Network analysis*. Beverly Hills, CA: Sage.
- Krackhardt, D. (1987). Cognitive social structures. *Social Networks*, 9, 109–134.
- Krackhardt, D. (1992). The strength of strong ties. In N. Nohria & R. G. Eccles (Eds.), *Networks and organizations: Structure, form and action* (pp. 216–239). Boston, MA: Harvard Business School Press.
- Krackhardt, D. (1994). Graph theoretical dimensions of informal organizations. In K. Carley & M. Prietula (Eds.), *Computational organizational theory* (pp. 89–111). Hillsdale, NJ: Lawrence Erlbaum.
- Kram, K. W. (1988). *Mentoring at work: Developmental relationships in organizational life*. New York, NY: University Press of America.
- Kraut, A. I. (1975). Predicting turnover of employees from measured job attitudes. *Organizational Behavior and Human Performance*, 13, 233–243.
- Kunda, Z. (2002). *Social cognition: Making sense of people* (2nd ed.). Cambridge, MA: The MIT Press.
- Lewin, A. Y., & Zwany, A. (1976). Peer nominations: A model, literature critique and a paradigm for research. *Personnel Psychology*, 29, 423–447.
- Leyens, J. P., & Fiske, S. T. (2002). Impression formation: From recitals to symphonie fantastique. In P. G. Devine, D. L. Hamilton, & T. M. Ostrom (Eds.), *Social cognition: Impact on social psychology* (pp. 39–75). San Diego, CA: Academic Press.
- London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. *Personnel Psychology*, 48, 803–839.
- Lord, R. G., Brown, D. J., Harvey, J. L., & Hall, R. J. (2001). Contextual constraints on prototype generation and their multilevel consequences for leadership perceptions. *Leadership Quarterly*, 12, 311–338.
- Lord, R. G., Foti, R. J., & Devader, C. L. (1984). A test of leadership categorization theory—Internal structure, information-processing, and leadership perceptions. *Organizational Behavior and Human Performance*, 34, 343–378.
- Louch, H. (2000). Personal network integration: Transitivity and homophily in strong-tie relations. *Social Networks*, 22, 45–64.

- Love, K. G. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias, and user reaction. *Journal of Applied Psychology*, 66, 451–457.
- Luria, G., & Torjman, A. (2009). Resources and coping with stressful events. *Journal of Organizational Behavior*, 30, 685–707.
- Marques, J. M., & Paez, D. (1994). The black sheep effect: Social categorization, rejection of ingroup deviates, and perception of group variability. *European Review of Social Psychology*, 5, 37–68.
- Mayfield, E. C. (1970). Management selection: Buddy nominations revisited. *Personnel Psychology*, 23, 377–391.
- Milliman, J. F., Zawacki, R. A., Norman, C., Powell, L., & Kirksey, J. (1994). Companies evaluate employees from all perspectives. *Personnel Journal*, 73, 99–103.
- Moreno, J. L. (1934). *Who shall survive? A new approach to the problem of human interrelations*. Washington, DC: Nervous and Mental Disease Pub. Co.
- Murphy, K. R., & Cleveland, J. N. (1991). Performance appraisal: An organizational perspective. Needham Heights, MA: Allyn and Baron.
- Murray, A. J. (1970). A sociometric approach to organizational analysis. *California Management Review*, 13, 59–67.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability on personality testing of personnel selection: The red herring. *Journal of Applied Psychology*, 81, 660–679.
- Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology*, 24, 1–24.
- Rank, O., Robins, G., & Pattison, P. (2010). Structural logic of intraorganizational networks. *Organization Science*, 21, 745–764.
- Reagans, R., & Zuckerman, E. (2001). Network, diversity and performance: The social capital of R&D units. *Organization Science*, 12, 502–517.
- Reilly, R. R., & Chao, G. T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 62, 1–62.
- Roadman, H. E. (1964). An industrial use of peer ratings. *Journal of Applied Psychology*, 48, 211–214.
- Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, 29, 173–191.
- Rodgers, F. (1992). A team approach to promotions at Rohr Inc. *Personnel Journal*, 71, 44–47.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances—Studies in internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Shamir, B., Zakay, E., Breinin, E., & Popper, M. (1998). Correlates of charismatic leader behavior in military units: Subordinates' attitudes, unit characteristics, and superiors' appraisals of leader performance. *Academy of Management Journal*, 41, 387–409.
- Shore, T. H., Shore, L. M., & Thornton, G. C., III. (1992). Construct validity of self- and peer evaluations of performance dimensions in an assessment center. *Journal of Applied Psychology*, 77, 42–54.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428.
- Simmel, G. (1955). *Conflict and the web of group affiliations*. Glencoe, IL: Free Press.
- Slater, P. E. (1955). Role differentiation in small groups. *American Sociological Review*, 20, 300–310.
- Smith, G. M. (1967). Usefulness of peer ratings of personality in educational research. *Educational and Psychological Measurement*, 27, 967–984.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Montgomery, CA: Brooks/Cole.
- Thomas, D. A. (1990). The impact of race on managers' experiences of developmental relationships (mentoring and sponsorship): An intra-organizational study. *Journal of Organizational Behavior*, 2, 479–492.
- Tichy, N. M., Tushman, M. L., & Fombrun, C. (1979). Social network analysis for organizations. *Academy of Management Review*, 4, 507–519.
- Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes* (Vol. 2, pp. 77–122). Greenwich, CT: JAI Press.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, UK and New York, NY: Basil Blackwell.
- Tushman, M., & Scanlan, T. J. (1981). Characteristics and external orientations of boundary spanning individuals. *Academy of Management Journal*, 24, 83–98.

- Waldman, D. A., Atwater, L. E., & Antonioni, D. (1998). Has 360 degree feedback gone amok? *Academy of Management Executive*, 12(2), 86–94.
- Wasserman, S., & Faust, K. (1994). *Social network analysis*. Cambridge, UK: Cambridge University Press.
- Wexley, K. N., & Klimoski, R. (1984). Performance appraisal: An update. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management* (pp. 35–79). Greenwich, CT: JAI Press.
- Willey, K., & Gardner, A. (2009). Improving self- and peer assessment processes with technology. *Campus-Wide Information Systems*, 26, 379–399.
- Yammarino, F. J., & Waldman, D. A. (1993). Performance in relation to job skill importance: A consideration of rater source. *Journal of Applied Psychology*, 78, 242–249.
- Yukl, G. (2002). *Leadership in organizations* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Yukl, G., Gordon, A., & Taber, T. (2002). A hierarchical taxonomy of leadership behavior: Integrating a half century of behavior research. *Journal of Leadership and Organization Studies*, 9, 15–32.
- Zammuto, R. F., London, M., & Rowland, K. M. (1982). Organization and rater differences in performance appraisals. *Personnel Psychology*, 35, 643–658.