

# UC Santa Barbara

## UC Santa Barbara Previously Published Works

**Title**

Jaccard-Spline index of structural proximity in contact networks

**Permalink**

<https://escholarship.org/uc/item/06x1r2fv>

**Journal**

Social Networks, 31(1)

**ISSN**

0378-8733

**Author**

Friedkin, Noah E

**Publication Date**

2009

**DOI**

10.1016/j.socnet.2008.10.002

Peer reviewed



Contents lists available at ScienceDirect

## Social Networks

journal homepage: [www.elsevier.com/locate/socnet](http://www.elsevier.com/locate/socnet)

## Jaccard-Spline index of structural proximity in contact networks

Noah E. Friedkin\*

Department of Sociology, University of California, Santa Barbara, CA 93106, United States

## ARTICLE INFO

**Keywords:**  
Contact  
Network  
Agreement  
Influence

## ABSTRACT

Network analysts are increasingly being called upon to apply their expertise to groups for which the only available or reliable data is a contact network. With no opportunity to gather additional data, the merits of such applications depend on empirical studies that validate the employment of structural constructs based on contact networks. Fortunately, we possess such studies in abundance. One of the strongest research traditions in social network analysis is the development of formal constructs that may be employed in analyses of networks. I suggest that greater insight into predictive success of network constructs may be acquired by addressing the following question: what features of the contact network in which a dyad is situated allow the prediction of other relations with an accuracy that validates the imputation of the latter given data on the former? In this article, I present findings on the structural contexts of dyads in contact networks and the relationship of these contexts with two fundamental forms of cohesive cognitive relations—accorded interpersonal influence and perceived interpersonal agreement. Based on these findings, I formalize a measure of structural proximity in contact networks with values that correspond to the conditional probabilities of these two forms of cohesive cognitive relations. The substantive settings of this analysis are policy groups with members who are embedded in contact structures based on regular interpersonal communication on policy issues and cognitive structures based on perceived interpersonal agreement and accorded interpersonal influence.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

Network analysis often proceeds on the basis of limited data on the social relations that are present in a group, and interpersonal contact is the relation that is most often available from persons' self-reported contacts, from observations and records of contacts, or from imputations based on self-reported, observed, or recorded instances of persons' co-presence in social foci where persons assemble and interact (Feld, 1981). Henceforth, the contact network is taken to be a connected undirected graph  $G$  with a vertex set  $V(G)$  of order  $n \geq 3$ .  $G$ 's corresponding symmetric adjacency matrix is denoted as  $\mathbf{G} = [g_{ij}]$ , where  $g_{ij} = 1$  indicates contact,  $g_{ij} = 0$  indicates non-contact, and  $g_{ij} = 0$  for all  $i$  and  $j$ . Ideally, based on the available contact network, the presence of other relations among group members, for which data is not available, may be imputed. That is, with structural constructs that are validated predictors of relations of interpersonal influence, agreement, trust, or affection, we may employ structural constructs as proxies for these relations or interpret structural formations in terms of the conditional probabilities of these relations, although the relations are unobserved. Such

employments dovetail with one of the strongest research traditions in social network analysis—the formal definition of constructs that draw on the information in  $G$ .

The employment of structural proxies and imputations are prevalent in the literature, especially for relations of interpersonal influence and agreement. Interpersonal influence is imputed from persons' point centralities in a contact network (Brass, 1984; Freeman, 1979; Friedkin, 1998). Relations of interpersonal agreement are imputed from persons' contacts (McPherson et al., 2001) and profiles of contacts (Burt, 1987; White et al., 1976). Interpersonal influence and agreement are imputed from persons' co-memberships in cliques, 2-plexes, 2-clans and other types of cohesive subgroups (Erickson, 1988; Mokken, 1979; Seidman and Foster, 1978; Friedkin, 2004). Such imputations are broadly justified by the large cumulative body of findings of social network research that has dealt with a variety of different types of groups in different cultures and times (e.g., Padgett and Ansell, 1993; Roethlisbrger and Dickson, 1939).

But it is also the case that the ingenuity of network analysts has generated a large set of structural constructs—alternative definitions of point centrality, social positions, and cohesive subgroups—all of which may be usefully employed as predictors of interpersonal influence and agreement. It appears difficult to be grossly misleading when we impute interpersonal agreement

\* Tel.: +1 805 893 2840; fax: +1 805 456 2075.  
E-mail address: [Friedkin@soc.ucsb.edu](mailto:Friedkin@soc.ucsb.edu).

and influence with any construct defined on a contact network, so long as the linkage is conceptually plausible, i.e., such imputation appears *robust* within the fuzzy domain of the plausible possibilities that have been entertained by network analysts. This result is both comforting and disconcerting. It is hard to go wrong, but it is not clear why. In such a situation, a natural hypothesis is that there are elementary constructs that are doing the predictive work and that higher-order constructs are drawing upon this work.

In this article, I focus on the relations of accorded interpersonal influence and perceived interpersonal agreement, reduce the prediction of these relations to the level of dyads, and rest the prediction on three elementary structural variables of the contact-network environment for each dyad—the presence or absence of direct contact ( $g_{ij}$ ), the number of mutual contacts ( $\sum_k g_{ik}g_{kj}$ ,  $k \neq i \neq j$ ), and the number of unique contacts ( $\sum_k (g_{ik} - g_{jk})^2$ ,  $k \neq i \neq j$ ). There is evidence that interpersonal visibility of role performance is improbable (near zero) in dyads with neither direct contact nor mutual contacts, and that the probability of such visibility increases with contact and the number of mutual contacts (Friedkin, 1986). It is plausible that the conditional probability of accorded interpersonal influence and perceived interpersonal agreement also depend on these local features of the contact structure in which a dyad is embedded—the presence/absence of direct contact and the number of mutual contacts, which are elementary indicators of co-orientation, as well as the number of unique contacts, which is an elementary indicator of idiosyncratic orientations.

Drawing on the above elementary features of the contact network  $G$ , I present empirical support for the following Jaccard-Spline index of structural proximity:

$$p_{ij} = \frac{g_{ij} + j_{ij}}{2} \quad (1)$$

where  $0 \leq j_{ij} \leq 1$  is the coefficient of Jaccard,  $j_{ij} = m_{ij}/(m_{ij} + u_{ij})$ ,  $m_{ij} = \sum_k g_{ik}g_{kj}$  ( $k \neq i \neq j$ ) is the number of mutual (matching) contacts, and  $u_{ij} = \sum_k (g_{ik} - g_{jk})^2$  ( $k \neq i \neq j$ ) is the number of unique (mismatching) contacts of  $i$  and  $j$  in  $G$ . Since  $G$  is a connected undirected graph of order  $n \geq 3$ ,  $m_{ij} + u_{ij} \neq 0$  for all  $i$  and  $j$  ( $i \neq j$ ).<sup>1</sup> I show that this proximity index has a strong association with the probability (relative frequency or density) of dyads in which interpersonal influence is accorded and interpersonal agreement is perceived.

The index is a simple but interesting formal model of structural proximity in  $G$ . It elevates the structural proximity of all contact dyads, as a class, above the proximity of all non-contact dyads, as a class. Within each of these classes, proximity is a function of the number of matching and mismatching contacts. An implication of the index is that the dyads within cohesive subgroups (e.g., 2-cliques, 2-clans, or 2-plexes in  $G$ ) may *markedly* differ in their probabilities of accorded interpersonal influence and perceived interpersonal agreement, depending on whether a dyad is a contact dyad or a non-contact dyad, and the dyad's ratio of mutual and unique contacts in  $G$ . In turn, the 2-cliques, 2-clans, or 2-plexes of  $G$  may markedly differ in their densities of accorded interpersonal influence and perceived interpersonal agreement. The same implications apply to social positions, i.e., subsets of persons with similar profiles of interpersonal contacts in  $G$ . Contact matters and so does the dyad's ratio of mutual and unique contacts in  $G$ .

Departures from the classic definition of a clique (Luce and Perry, 1949) allow non-contact dyads within cliques. The Jaccard-Spline index situates accorded interpersonal influence and perceived

interpersonal agreement in 2-cliques (Alba, 1973).<sup>2</sup> The present empirical evidence indicates that the conditional probability of each of these relations is near zero given both  $g_{ij} = 0$  (non-contact) and  $m_{ij} = 0$  (no mutual contacts). However, while non-zero probabilities of accorded interpersonal influence and perceived interpersonal agreement occur in the 2-cliques of the contact network, these probabilities need not be substantial. A non-contact dyad of a 2-clique must have at least one mutual contact. The conditional probability of accorded interpersonal influence and perceived interpersonal agreement in a non-contact dyad of a 2-clique depends its ratio of unique and mutual contacts  $u_{ij}/m_{ij}$  in  $G$ . For a non-contact dyad within a 2-clique of  $G$ ,  $p_{ij} = 0.50(1 + u_{ij}/m_{ij})^{-1}$ . The larger the ratio  $u_{ij}/m_{ij}$ , the lower the structural proximity of  $i$  and  $j$ , and the lower the probabilities of accorded interpersonal influence and perceived interpersonal agreement for a non-contact dyad. Proximity approaches  $p_{ij} = 0$  when the number of the non-contact dyad's unique contacts swamp the number of the dyad's mutual contacts, and it approaches  $p_{ij} = 0.50$  when the number of the non-contact dyad's mutual contacts swamp the number of the dyad's unique contacts. The maximum proximity of a non-contact dyad is attained with structural equivalence, i.e., when  $g_{ij} = 0$  and  $u_{ij} = 0$ . Here, a non-contact dyad's number of unique contacts is contextualized by the number of the dyad's mutual contacts, and vice versa.

Contact matters. The index attains its maximum value ( $p_{ij} = 1$ ) with contact *and* structural equivalence, i.e., when  $g_{ij} = 1$  and  $u_{ij} = 0$ . In  $G$ , the absence of unique contacts indicates that  $i$  and  $j$  must have some number of mutual contacts greater than zero; with  $u_{ij} = 0$  the index is at its maximum for *all* values of  $m_{ij} \geq 1$ . An implication, which I am unable to evaluate with the present data, is that if  $G$  is a Luce-Perry clique, i.e.,  $g_{ij} = 1$  and  $u_{ij} = 0$  for all  $i$  and  $j$  in  $G$ , then the densities of accorded interpersonal influence and perceived interpersonal agreement will be substantial, regardless of order of  $G$ . In general, for a contact dyad of a 2-clique of  $G$  ( $g_{ij} = 1$ ,  $m_{ij} \geq 1$ ), the ratio  $u_{ij}/m_{ij}$  matters: in such a case,  $p_{ij} = 0.50 + 0.50(1 + u_{ij}/m_{ij})^{-1}$ . The larger  $u_{ij}/m_{ij}$ , the lower the structural proximity of the contact dyad, and the lower the probabilities of accorded interpersonal influence and perceived interpersonal agreement. Proximity approaches  $p_{ij} = 0.50$  when the number of the contact dyad's unique contacts swamp the number of the dyad's mutual contacts, and it approaches  $p_{ij} = 1$  when the number of the contact dyad's mutual contacts swamp the number of the dyad's unique contacts. Again, the dyad's number of mutual contacts is structurally contextualized by the dyad's number of unique contacts, and vice versa. Burt (1999) emphasizes that "equivalence corrects cohesion," but here it also appears that cohesion corrects equivalence.

The "knot" of the Jaccard-Spline index is  $p_{ij} = 0.50$ , which is minimum value of the index for contact dyads and the maximum value of the index for non-contact dyads. In  $G$ , a contact dyad at the "knot" is a local bridge, i.e., a contact with no mutual contacts with at least one unique contact in  $G$  (Granovetter, 1973). In  $G$ , a non-contact dyad at the "knot" arises in the absence of unique contacts. Since in  $G$ , a non-contact dyad without unique contacts must have at least one mutual contact, the proximity of such a structurally equivalent non-contact dyad is  $p_{ij} = 0.50$ , regardless of the number of their mutual contacts. Thus, the proximity of a non-contact dyad cannot exceed  $p_{ij} = 0.50$  (the minimum value of the index for a contact dyad), and the proximity of a contact dyad cannot be lower than  $p_{ij} = 0.50$  (the maximum value of the index for a non-contact dyad). Burt (1999) emphasizes that "equivalence extends cohesion," but

<sup>1</sup> Since  $G$  is a connected undirected network of order  $n \geq 3$ , every  $i, j$  dyad in  $G$  has at least one contact that is either a unique or mutual contact of the dyad.

<sup>2</sup> A 2-clique of  $G$  is a maximal subgraph in which every pair of vertices is connected by a path of length 2 or less in  $G$ , i.e., every pair is either in contact or has at least one mutual contact in  $G$ .

here it does so to a limited extent. It is only among contact dyads that proximities may attain high levels.

In short, the present article contributes a simple index of structural proximity in contact networks that entails important linkages to the comparative literature on structural cohesion and equivalence (e.g., Burt, 1978, 1987; Friedkin, 1984; Meyer, 1994; Mizruchi, 1993). The present article moves toward a possible synthesis of the competitive hypotheses that have been entertained on the relative merits of these two structural approaches. Contact alone provides a powerful basis for predicting the probability of accorded interpersonal influence and perceived interpersonal agreement. However, within the classes of contact and non-contact dyads, the number of mutual and unique contacts discriminate conditions that are related to the probabilities of these cognitive relations. The Jaccard-Spline index is a formalization of the observed pattern of findings presented in this article. However, it not a synthesis that strongly favors the structural equivalence approach, where its proponents have emphasized that contact is *irrelevant* to the imputation of relations of interpersonal influence and agreement (e.g., Burt, 1987). The hypothesis that the members of a structurally equivalent position are an important reference group for the position's occupants, regardless of the occupants' contacts with one another, becomes difficult to sustain when such a position does not entail high densities of accorded interpersonal influences or perceived interpersonal agreements. The absence of unique contacts (structural equivalence) corresponds to the maximum values of the index *within* the classes of contact and non-contact dyads; however, the probability of accorded interpersonal influence and perceived interpersonal agreement is substantially higher in the former than in the latter class. Moreover, an emphasis on mismatches and de-emphasis of matches is troublesome when one measure provides an informative structural context for the other. My hope is that the properties of the Jaccard-Spline index, and the empirical analyses to be presented, will serve to shift the debate on structural cohesion and equivalence toward a more integrative approach in which hegemonic assertions in favor of one or the other are discarded.

The proximity matrix of the Jaccard-Spline index,  $\mathbf{P} = [p_{ij}]$ , possesses the following standard properties: non-negativity,  $p_{ij} \geq 0$  for all  $i$  and  $j$ , symmetry,  $p_{ij} = p_{ji}$  for all  $i \neq j$ , and identity,  $(p_{ii} = p_{jj}) > p_{ij}$ , for all  $i \neq j$  (Batagelj and Bren, 1995, p. 74). The third property requires a scalar constant on the main diagonal of  $\mathbf{P}$  that is larger than any of the off-diagonal values of  $\mathbf{P}$ . Based on the present empirical findings, it also is a measure on an *absolute* scale with a minimum  $p_{ij} = 0$  that indicates the near certain absence of accorded interpersonal influence and perceived interpersonal agreement. My evidence indicates that the probabilities of each cognitive relation are approximately a linear function of increasingly index values. For  $p_{ij} = 0.50$  these probabilities are near 0.50. The probabilities of each relation become substantial for index values  $p_{ij} > 0.75$ , i.e., for contact dyads with a number of mutual contacts that exceeds the number of unique contacts ( $m_{ij} > u_{ij}$ ). Hence, the available empirical evidence roughly justifies an imputation that the relation of accorded interpersonal influence or perceived interpersonal agreement is  $x$  times as likely in one dyad than in another, based on the index values for the two dyads. For proximity indices without a meaningful zero-point, proximity values may either indicate trivial or substantial differences among dyads, and such comparative conclusions may be unreliable. In turn, absent such an index, higher-order constructs, such as cliques and social positions, may indicate trivial or substantial levels of influence or agreement, and such comparative conclusions also may be unreliable.

The findings to be presented deal with the relative frequency (density) of accorded interpersonal influence and perceived interpersonal agreement in dyads conditioned on features of the local

contact network. The imputation of a *specific instance* of accorded influence or perceived agreement (i.e., a relation from a particular  $i$  to a particular  $j$ ) is a reliable imputation near the *extremes* of the index. It appears that a reliable imputation of a specific instance of accorded interpersonal influence or perceived interpersonal agreement may be made for index values exceeding 0.75; i.e., for a contact dyad with a number of mutual contacts that exceeds the number of unique contacts.

## 2. Data and measures

For this analysis, I employ data on six policy groups in which both formal and informal bases of regular interpersonal contact are present, and that provide two measures of fundamental forms of cohesive interpersonal cognitions—accorded interpersonal influence and perceived interpersonal agreement on policy group issues. The contact networks of these groups are analyzed as undirected graphs and their cognitive networks as digraphs. For the definition of the former, a report of periodic communication by either  $i$  or  $j$  suffices to establish the occurrence of contact between the two. In principle, an outside observer might observe or verify whether or not two persons have been in contact or not. In contrast, cognitive relations may only be directly ascertained from the *source*, who is the holder of the cognition, and it refers to a source-target relation that is in the mind of the source. My inquiry deals with the conditional density of cohesive interpersonal cognitions under different structural conditions of the contact network in which two persons are embedded. For an alternative approach, which treats  $i$ 's report of regular contact with  $j$  as evidence of  $i$ 's cognitive orientation toward  $j$ , i.e., as a digraph, rather than as evidence of contact per se, see Friedkin (1998).

### 2.1. Six policy groups

Policy groups occur wherever there are persons who have authority to make decisions in a particular policy domain; however, the formal decision makers in these groups are not necessarily accorded influence by all other persons in the group or perceived as being in general agreement with them. More or less stable structures of interpersonal contact in policy groups are constructed and maintained by the activities of persons who must or might make decisions in a particular domain, and by the activities of persons whose advice is sought by, or who attempt to influence the decisions of, these decision makers.

The particular substantive focus of my investigation is school board policy groups in six US public school districts. These groups consist of the school board members of a school district (usually seven elected residents of a community who have formal authority over all school district issues and personnel) and those persons in the school district (administrators, faculty, and community residents) who have some influence on school board decisions. These groups have a formal structure of authority in which policy decisions are made by a small set of decision makers, but the groups also are highly political. School Board members are elected officials, and the School District Superintendent is appointed by the Board. Without the support of the Board, the position of the Superintendent is put at risk. Without community understanding and support for the policies recommended and enacted by the Board and Superintendent, the positions of both are put at risk. School board deliberations are open to public scrutiny and influence so that the members of the local community, along with the district's administrators, may become involved in the board's decision making. Hence, policies are usually formulated and implemented via discussion and consultation.

**Table 1**  
Six policy groups.

Group	<i>n</i>	Survey respondents	Ordered dyads	Ordered respondent dyads	Unordered respondent dyads
A	42	37*	1722	1332	666
B	42	39	1722	1482	741
C	60	59	3540	3422	1711
D	61	55	3660	2970	1485
E	59	57	3422	3192	1596
F	67	66	4422	4290	2145
Totals	331	313	18488	16688	8344

Note\*: The non-respondents of Group A include two isolates without contacts. No such isolates occurred in the remaining groups. With the isolates of Group A removed, all of the contact networks *G* of Groups A–F are connected.

## 2.2. Definition of the group

In each district, the policy group was defined by a snowball procedure. This procedure combined positional, reputational, and behavioral selection criteria. The initial sample included the members of the district's school board, the district's superintendent, the district's school principals, and all others identified in local newspapers or board minutes as currently or recently active in board meetings, district elections, or other efforts to affect board decisions. These individuals nominated persons on the district staff and in the community who were known or reputed to be currently or recently active in attempts to influence school board decisions. These nominees were asked for further nominations according to the same criteria. To guard against idiosyncratic nominations, persons were added to the group only if they were mentioned at least twice. The nominating procedure continued until dual nominations no longer occurred. The policy groups that emerged contained between 42 and 67 members.

## 2.3. Data

Questionnaires were administered to all the persons who had been identified as policy group members. The questionnaire contained items concerned with group members' interpersonal relations, attitudes and opinions on various topics, and socioeconomic characteristics. For the relational data, the questionnaire contained a roster of the group members on which a respondent checked the names of those members with whom a particular type of relation existed. The response rates to the questionnaires were 88% in Group A ( $n=42$ ), 93% in Group B ( $n=42$ ), 98% in Group C ( $n=60$ ), 90% in Group D ( $n=61$ ), 97% in Group E ( $n=59$ ), and 99% in Group F ( $n=67$ ); see Table 1.

I draw on three binary social relations from this survey: *interpersonal contact* ("Is this a person with whom you frequently discuss matters having to do with the . . . schools?"), *accorded interpersonal influence* ("Is this a person who you would say probably had some influence on your own opinions on school related matters during the last year or so?"), and *perceived agreement* ("When there are differences of opinion about school district matters, is this a person who is usually on the same side of these issues as yourself?"). The two binary cognitive relations are analyzed without transformation. The adjacency matrix of reported contacts,  $C=[c_{ij}]$ , is transformed into  $G=[g_{ij}]$  as follows:  $g_{ij}=\max(c_{ij}, c_{ji})$  for all  $i$  and  $j$ , and  $g_{ii}=0$  for all  $i$ . In five of the six groups (Groups B–F), the contact network is connected. Group A contains two isolates, and the connected subnet of the remaining members is analyzed.

Contact with non-respondents may be reported by respondents. Assuming that the incidence of false evidence of non-contact is slight among dyads with at least one respondent, false non-contact (i.e., contact dyads with no evidence of contact) will be concentrated in those dyads with two non-respondents. In the

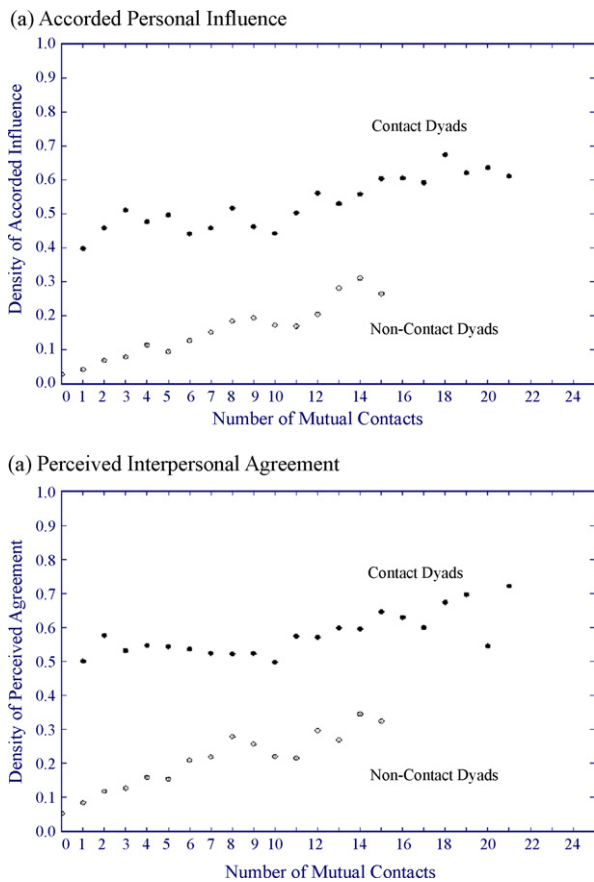
present data, the numbers of non-respondents in each group is small (Table 1). Among the respondent dyads in which at least one member reported contact, the percentage of joint (reciprocated) reports of contact is surprisingly low (36.4%). This modest level of joint-contact reports is probably due to a combination of factors including false positive reports, false negative reports, and differences in respondents' interpretation of the phrase "frequently discuss." Reporting that a particular person is someone with whom a respondent has frequent discussions also may be affected by the cognition orientation of the source toward the target as a salient or significant other. I rely on the assumption that *periodic* discussions of policy group matters have occurred in dyads with at least one member who has reported the occurrence of frequent discussion.

## 2.4. Findings

Fig. 1 displays the conditional density of ordered pairs in which  $i$  accords personal influence to  $j$  (Fig. 1a) and in which  $i$  reports agreement with  $j$  on policy-related issues (Fig. 1b). Each of the displayed densities is based on a subset of the 16,688 ordered respondent dyads (Table 1) defined by the presence (or absence) of contact and the number of a dyad's mutual contacts. For a subset with 30 or more replicate dyads (e.g., 30 or more non-contact dyads with four mutual contacts), the conditional density of accorded interpersonal influence is the number of dyads in the subset in which  $i$  accords influence to  $j$  divided by the number of dyads in the subset (Fig. 1a), and the conditional density of perceived interpersonal agreement is the number of dyads in the subset in which  $i$  perceives  $j$  as being in general agreement with  $i$  divided by the number of dyads in the subset (Fig. 1b).

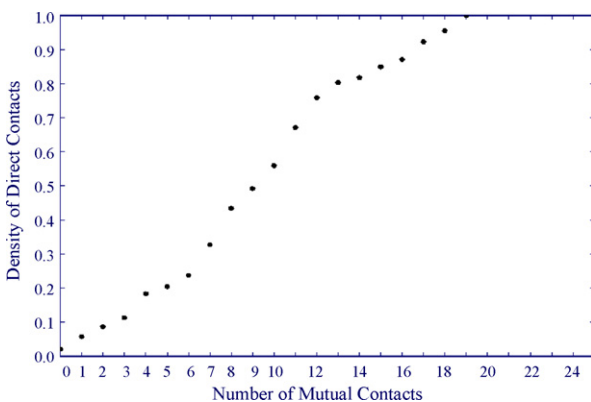
There are three noteworthy features of the plots in Fig. 1. The densities of these interpersonal cognitions are substantially greater among contact dyads than among non-contact dyads. The densities of these interpersonal cognitions increase with the number of mutual contacts. The densities of these interpersonal cognitions among non-contact dyads, with a large number of mutual contacts, are lower than the densities of such cognitions among contact dyads with few mutual contacts. The first feature is not a surprise, nor is the second. It is the form of the associations that is noteworthy. The densities of these interpersonal cognitions increase from near zero, in the absence of contact and mutual contacts, to substantial densities, in contexts entailing contact and numerous mutual contacts. The third feature has important implications. The densities of these interpersonal cognitions among non-contact dyads has a ceiling, regardless of the number of mutual contacts, that is below the floor of local bridges, i.e., contacts with no mutual contacts.

A hidden feature of these plots is that a non-contact dyad with numerous mutual contacts is a rare event. Fig. 2 shows that as the number of mutual contacts increases, so does the density of direct contact. While violations of transitive closure in contact networks are numerous, the probability of such closure increases with the



**Fig. 1.** Accorded personal influence and perceived interpersonal agreement. *Note:* These plots are based on 16,688 ordered respondent dyads. Only densities based on 30 or more dyads are displayed. The open circles present the densities of perceived interpersonal agreement for non-contacts dyads, and the solid circles the densities of perceived interpersonal agreement for contact dyads.

number of mutual contacts. In the adjacency matrix of the contact network, the “if–then” logic of transitivity arises when  $g_{ik} = 1$  and  $g_{kj} = 1 (i \neq j \neq k)$ ; transitivity is satisfied when  $g_{ij} = 1$  and violated when  $g_{ij} = 0$ . If the probability of transitive closure were high, then the curve in Fig. 2 would rise dramatically over a small range of mutual contacts. If it were never violated, then the density of contact dyads among dyads with one mutual contact would be 1. The



**Fig. 2.** Transitive closure in contact networks. *Note:* In the adjacency matrix of the contact network  $G$ , transitive closure occurs when  $g_{ij} = 1 (i \neq j)$  given  $g_{ik} = 1$  and  $g_{jk} = 1$  for some  $k \neq i, j$ . This plot is based on 8,488 unordered dyads. Only densities based on 30 or more dyads are displayed. The solid circles present the densities of contact dyads.

consequence of the structural tendency to transitive closure is a *gradient* of diminishing probability of non-contact with an increase in the number of mutual contacts. With the satisfaction of transitive closure in the contact network, a dyad transitions into a higher state of cognitive cohesion (i.e., a state entailing a higher probability of accorded interpersonal influence and perceived interpersonal agreement) than the cognitive cohesion of non-contact dyads with larger numbers of mutual contacts.

A linear combination logic is misleading with such data, and a *spline* approach (Ahlberg et al., 1967) provides a more accurate representation of the organization of relational cohesion. It is evident that the probability of accorded interpersonal influence and perceived interpersonal agreement in non-contact dyads increases with the number of mutual contacts, but one should not extend the trend to include numbers of mutual contacts among the non-contact dyads that rarely, or do not, empirically exist. Extending the curves, shown in Fig. 1 for the non-contact dyads into the region of exceedingly rare conditions—non-contact in combination with numerous mutual contacts—implies that equivalent probabilities of these interpersonal cognitions may occur for contact and non-contact dyads given a sufficient number of mutual contacts for the latter. But it does not appear that mutual ties may substitute for the absence of direct contact, and it seems that a non-contact dyad with a large number of mutual contacts is a structural anomaly. A spline measure, which dovetails the two curves of Fig. 1, provides a more accurate description of the structural conditions of cohesive interpersonal cognitions than a linear combination without restrictions on anomalous combinations of values on the variables.

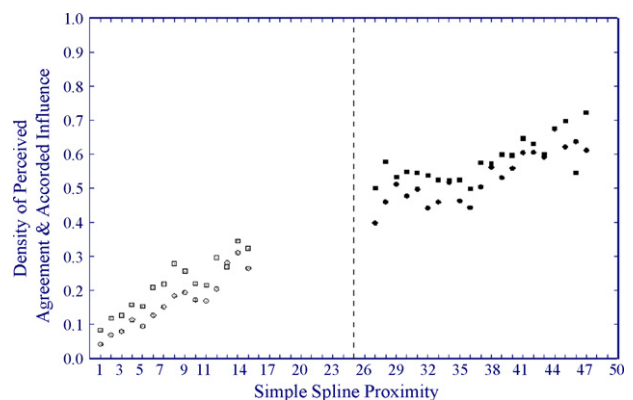
### 2.5. Jaccard-Spline index

In this section, the pattern described in Fig. 1 is formalized with the Jaccard-Spline index of structural proximity in a contact network. As a segue to this index, consider Fig. 3. Here the data are plotted against the following simple rank-order spline:

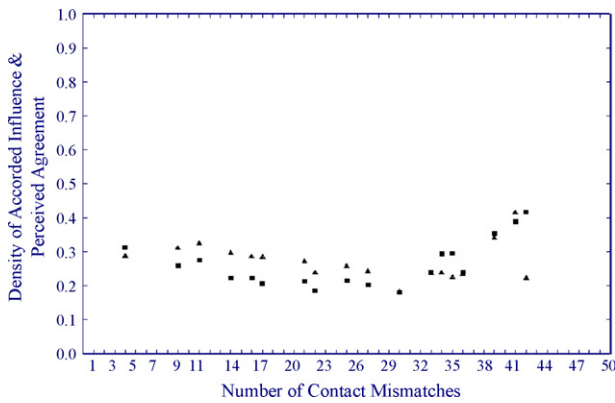
$$m_{ij} + g_{ij}(1 + h) \tag{2}$$

where  $g_{ij} = 1$  if  $i$  and  $j$  are a contact dyad,  $g_{ij} = 0$  if  $i$  and  $j$  are a non-contact dyad,  $m_{ij}$  is the number of mutual contacts of  $i$  and  $j$ , and  $h$  (25 in these data) is the observed maximum number of mutual contacts among the non-contact dyads.

The plotted densities in Fig. 3 are *exactly* those presented in Fig. 1, with the contact-based points shifted to the right. The den-



**Fig. 3.** Simple Spline Proximity in Contact Networks. *Note:* This plot is based on 16,688 ordered respondent dyads. Only densities based on 30 or more dyads are displayed. The vertical dashed line marks the maximum number of mutual contacts (25) observed among the non-contact dyads in  $G$ . The open circles and squares present the densities for non-contacts dyads, and the solid circles and squares the densities for contact dyads. Circles present the densities of accorded interpersonal influence, and squares perceived interpersonal agreement.



**Fig. 4.** Cohesive interpersonal cognitions and contact profile dissimilarity. *Note:* This plot is based on 16,688 ordered respondent dyads. Only densities based on 30 or more dyads are displayed. The squares present the densities of accorded influence, and the triangles the densities of perceived agreement.

sities increase from near 0 to near 0.80 as a function of contact and mutual contacts. The “vacuous gap” observed in the plot for proximity values 16–25 (i.e., non-contact dyads with 16–25 mutual contacts) is indicative of the rarity of dyads without direct contact and numbers of mutual contacts in this range. In analyses not presented, I find that that perceived disagreement (“Is this a person who is usually on a different side of these issues from yourself?”), and perceived friendship (“Is this person a close personal friend?”), are evenly distributed across the measure. In these data, perceived friendship is rare in non-contact dyads, regardless of the number of mutual contacts; and while perceived friendship is more likely among contact dyads, it is not associated with the number of mutual contacts.

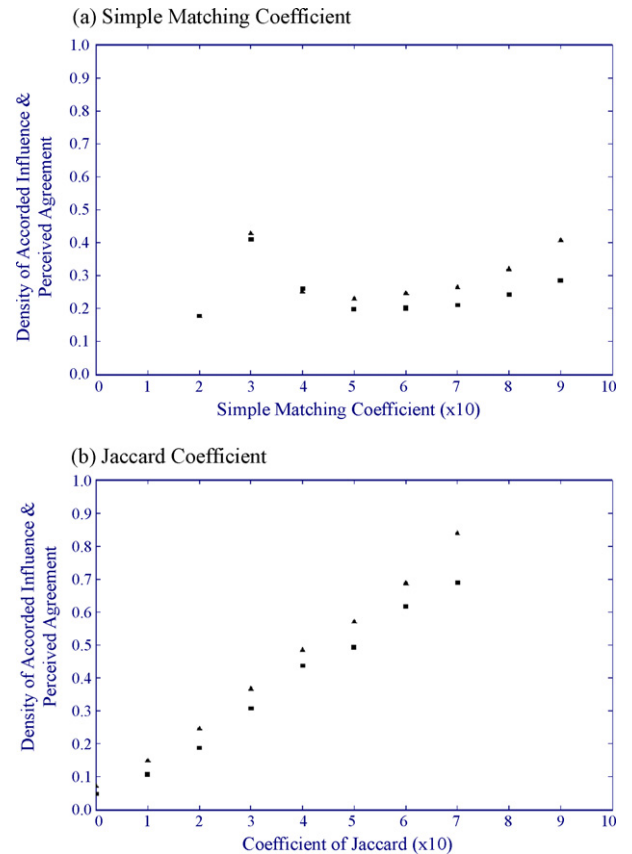
The unusual features of this measure are highlighted with a comparison to an approach based on the profile dissimilarity of two persons in the contact network

$$d_{ij}^2 = \sum_k (g_{ik} - g_{jk})^2 \quad k \neq i \neq j \quad (3)$$

Two persons are structurally equivalent in  $G$  when  $d_{ij}^2 = 0$ , and their extent of profile dissimilarity is indexed by the value of  $d_{ij}^2$  (which is equivalent to the number of unique contacts  $u_{ij}$ ). The presence or absence of contact between  $i$  and  $j$  does not affect the value of this measure.<sup>3</sup> A fundamental underpinning of the conceptualization of structural equivalence is that two persons with similar profiles of contacts, regardless of whether they are in contact with each other, are likely to be similar in other respects (e.g., are likely to perceive a structurally equivalent other as in agreement) and serve as referents for their attitudes and behaviors (e.g., to accord influence to a structurally equivalent other). The assumed underpinning of meaning is not supported in these data. Fig. 4 plots the densities of accorded influence and perceived agreement for each value of  $d_{ij}^2$  with 30 or more dyads as the basis of the density value. The solid squares present the densities of accorded influence, and the solid triangles present the densities of perceived agreement. This measure of structural equivalence does not discriminate dyads in which the two cognitive relations are more or less likely.

Fig. 5 displays the results for two standard widely applied measures of proximity—the Simple Matching and Jaccard coefficients. The Jaccard coefficient has been previously defined. In  $G$ , the Simple Matching coefficient is  $(m_{ij} + n_{ij}) / (n - 2)$ , where  $n_{ij}$  is number of

<sup>3</sup> This measure is sometimes elaborated as follows:  $d_{ij}^2 = (g_{ij} - g_{ji})^2 + \sum_k (g_{ik} - g_{jk})^2$ . In  $G$ , such elaboration is moot.

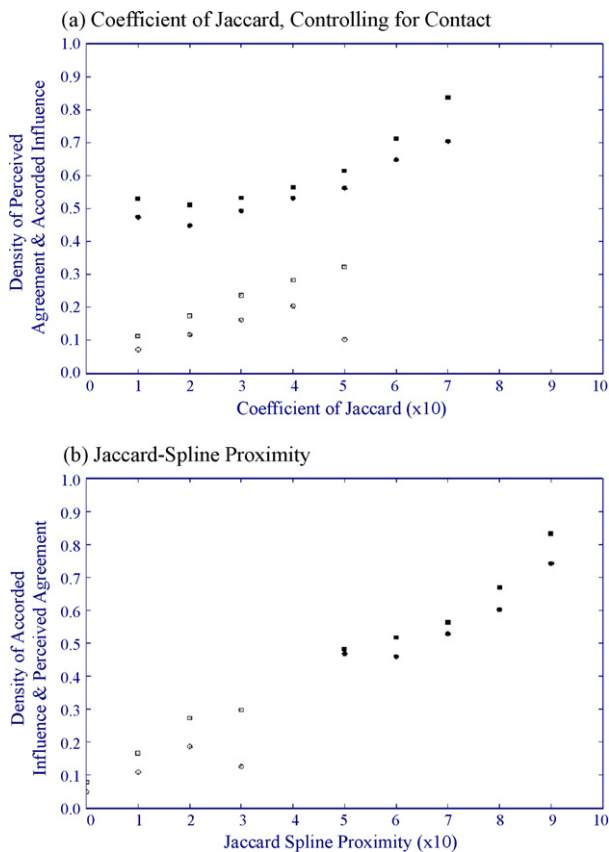


**Fig. 5.** Simple matching and Jaccard coefficients of proximity. *Note:* This plot is based on 16,688 ordered respondent dyads. Only densities based on 30 or more dyads are displayed. For this plot, the proximity values have been multiplied by 10 and rounded to nearest integer. The squares present the densities of accorded influence, and the triangles the densities of perceived agreement.

a dyad’s null matches, i.e., the number of persons  $k \neq \{i, j\}$  in  $G$  with whom neither  $i$  nor  $j$  are in contact. For these plots, the proximity values have been multiplied by 10 and rounded to nearest integer. As in the previous plots, only those densities of accorded influence and perceived agreement are displayed that are based on 30 or more dyads.

The Simple Matching coefficient does not discriminate dyads in which the interpersonal cognitions are more or less likely, while the Jaccard coefficient does. The former measure includes persons as matches with whom neither  $i$  nor  $j$  are in direct contact; the latter measure does not. The number of group members  $k \neq \{i, j\}$  with whom neither  $i$  nor  $j$  have direct contact appear irrelevant to the occurrence of these interpersonal cognitions. No doubt, the joint absence of contact with a third person may be a meaningful occurrence in some cases; however, the vast majority of such null matches arise on other less meaningful grounds. For instance, in a sufficiently large network, the number of non-contact dyads will be large and so will the number of persons with whom both  $i$  and  $j$  have no contact. The Jaccard measure was developed to deal with the swamping effect of such null matches. In its present application, the Jaccard measure draws its strength from the powerful tendency toward transitive closure, described previously in Fig. 2. As the number of mutual contacts increases, so does the density of direct contact and, in turn, the density of accorded influence and perceived agreement.

The Jaccard measure is widely available in data analysis software programs, and it seems worthwhile to base a spline measure of structural proximity upon it. In contrast to the Simple Rank-



**Fig. 6.** Coefficient of Jaccard and Jaccard-Spline proximities. *Note:* This plot is based on 16,688 ordered respondent dyads. Only densities based on 30 or more dyads are displayed. For this plot, the proximity values have been multiplied by 10 and rounded to nearest integer. The open circles and squares present the densities for non-contacts dyads, and the solid circles and squares the densities for contact dyads. Circles present the densities of accorded interpersonal influence, and squares perceived interpersonal agreement.

Order Spline (2), which only incorporates information on contact and mutual contacts, the Jaccard-Spline (1) also incorporates information on a dyad's unique contacts and, in turn, presents suggestive linkages to the literature on structural cohesion and equivalence. Fig. 6a shows that dyads with similar Jaccard proximities differ in the density of these interpersonal cognitions depending on whether contact is present in the dyad or not. High densities of these interpersonal cognitions are reached only among contact dyads. Fig. 6b displays the densities of accorded influence and perceived agreement for Jaccard-Spline index values that have been multiplied by 10 and rounded to the nearest integer.

Table 2 compares the Jaccard-Spline measure and the profile dissimilarity measure (3). A logistic regression of accorded interpersonal influence and perceived interpersonal agreement on each measure is presented along with selected predicted probabilities. The results are consistent with Figs. 4 and 6. Profile dissimilarity based strictly on the number of unique (mismatching) contacts is a relatively indiscriminating measure with respect to the occurrence of these two cognitive relations.

### 2.6. Detailed predictions

In the present data, there are no instances of strict structural equivalence, i.e., dyads with one or more mutual contacts and no unique contacts  $\{m_{ij} \geq 1, u_{ij} = 0\}$ . The available dyads either have no mutual contacts and at least one unique contact  $\{m_{ij} = 0, u_{ij} \geq 1\}$ , or they have at least one mutual contact and at least one unique con-

tact  $\{m_{ij} \geq 1, u_{ij} \geq 1\}$ . Among the respondent dyads, 6.6% are dyads of the first type  $\{m_{ij} = 0, u_{ij} \geq 1\}$  and 93.4% are dyads of the second type  $\{m_{ij} \geq 1, u_{ij} \geq 1\}$ . Recall that a dyad in  $G$  must have at least one mutual or at least one unique contact, i.e.,  $m_{ij} + u_{ij} \geq 1$  for all  $i$  and  $j$ .

For dyads of the first type  $\{m_{ij} = 0, u_{ij} \geq 1\}$ , the Jaccard-Spline index predicts that the number unique contacts has *no effect* on structural proximity:  $p_{ij} = 0.50g_{ij}$ . This prediction is supported by the results presented in Table 3, where accorded interpersonal influence and perceived interpersonal agreement are regressed on the number of unique contacts. Among non-contact dyads with no mutual contacts, the number of unique contacts has no significant effect ( $p > 0.10$ ), and the intercept locates the probabilities of these cognitive relations near zero.<sup>4</sup> Among contact dyads with no mutual contacts (i.e., local bridges), neither the intercept nor the effect of unique contacts significantly differ from zero ( $p > 0.10$ ); hence, the probabilities of these cognitive relations among local bridges are located near 0.50.<sup>5</sup> The number of available local bridges is small; hence, the more compelling of these two supports is that obtained for the non-contact dyads. Non-contact dyads with no mutual contacts entail low probabilities of accorded interpersonal influence and perceived interpersonal agreement regardless of their extent of structural equivalence.

For dyads of the second type  $\{m_{ij} \geq 1, u_{ij} \geq 1\}$ , which comprise the bulk of the available dyads, the Jaccard-Spline index predicts that within each class of dyads (contact and non-contact dyads) the probability of accorded interpersonal influence and perceived interpersonal agreement (a) increases with the number of mutual contacts, controlling for the number of unique contacts, and (b) declines with the number of unique contacts, controlling for the number of mutual contacts. Although the number of unique contacts alone clearly does not provide adequate discrimination of the conditions of accorded interpersonal influence and perceived interpersonal agreement, the number of a dyad's unique contacts should be negatively associated with the probability of these interpersonal cognitions controlling for the occurrence of contact and the number of mutual contacts. Table 4 evaluates whether this expected conditional association is present. It is supported by three of the four tests.

### 3. Discussion

Contact and mutual contacts are direct indicators of co-orientation and are powerful predictors of accorded interpersonal influence and perceived interpersonal agreement. Idiosyncratic unique contacts discriminate dyads that are otherwise similar. The present investigation suggests that contact should not be treated merely as an instance of co-orientation that is folded into the number of mutual contacts, but indicates a powerful condition that corresponds to markedly elevated levels of cognitive cohesion. It appears that non-contact dyads with large numbers of mutual contacts or small numbers of unique contacts on average entail *lower* probabilities of accorded interpersonal influence and perceived interpersonal agreement than do contact dyads with few mutual contacts or large numbers of unique contacts.

The debate within the field on the relative merits of structural cohesion and equivalence as explanatory constructs might be productively shifted toward approaches that pursue a synthe-

<sup>4</sup> Among the 1082 non-contact dyads with no mutual contacts, the empirical density of accorded interpersonal influence and perceived interpersonal agreement is 0.027 and 0.052, respectively.

<sup>5</sup> Among the 22 contact dyads with no mutual contacts, the empirical density of accorded interpersonal influence and perceived interpersonal agreement is 0.50 and 0.273, respectively.



**Table 2**

Comparison of the Jaccard-Spline and profile dissimilarity measures: logistic regression of accorded interpersonal influence and perceived interpersonal agreement on each measure ( $N = 16,688$ ; standard errors in parentheses).

	Accorded influence	Perceived agreement
<b>(a) Jaccard-Spline index</b>		
Constant	-2.58747*** (0.0359)	-2.07275*** (0.0300)
Jaccard-Spline	3.97932*** (0.0720)	3.54033*** (0.0661)
Selected predicted probabilities		
Jack-Spline = 0	0.069	0.112
Jack-Spline = 0.25	0.169	0.234
Jack-Spline = 0.50	0.355	0.425
Jack-Spline = 0.75	0.598	0.642
Jack-Spline = 1	0.801	0.813
<b>(b) Unique contacts (Profile Dissimilarity)</b>		
Constant	-1.21156*** (0.0530)	-0.58851*** (0.0491)
Unique contacts	-0.00337 (0.0024)	-0.02048*** (0.0023)
Selected predicted probabilities		
Unique contacts = 0	0.229	0.357
Unique contacts = 5	0.226	0.334
Unique contacts = 10	0.224	0.311
Unique contacts = 15	0.221	0.290
Unique contacts = 20	0.218	0.269

\*\*\*  $p < .001$ .

sis of the two approaches. The conclusion appears warranted that the structural indicators of co-orientation and idiosyncratic orientation, which each approach separately emphasizes, should be combined and that any hegemonic assertion of the superiority of one approach over the other is misplaced. A selective emphasis on mismatches or matches is problematic. A discounting of contact appears entirely unjustified. The Jaccard-Spline index presents a useful formal perspective in which both structural constructs are implicated and serve together to predict interpersonal influences and agreements. My hope is that the index will encourage the development of a new class of structural measures, and I propose it as a stepping stone to such developments. These developments should

**Table 3**

Dyads with no mutual contacts: logistic regression of accorded interpersonal influence and perceived interpersonal agreement on profile dissimilarity (standard errors in parentheses).

	Accorded influence	Perceived agreement
Non-contact dyads ( $N = 1,082$ )		
Constant	-4.21593*** (0.5020)	-2.39761*** (0.3461)
Unique contacts	0.03621 (0.0257)	-0.03251 (0.0211)
Contact dyads ( $N = 22$ )		
Constant	-0.26559 (1.0264)	-1.74192 (1.1894)
Unique contacts	0.01359 (0.0478)	0.03770 (0.0523)

\*\*\*  $p < .001$ .

**Table 4**

Dyads with at least one mutual contact and at least one unique contact: logistic regression of accorded interpersonal influence and perceived interpersonal agreement on the number of mutual and unique contacts (standard errors in parentheses).

	Accorded influence	Perceived agreement
Non-contact dyads ( $N = 10,742$ )		
Constant	-3.06076*** (0.1068)	-1.76459*** (0.0839)
Mutual contacts	0.13759*** (0.0092)	0.13068*** (0.0081)
Unique contacts	0.00936* (0.0045)	-0.02573*** (0.0038)
Contact dyads ( $N = 4,842$ )		
Constant	-0.15259† (0.0909)	0.47701*** (0.0919)
Mutual contacts	0.04577*** (0.0057)	0.03421*** (0.0058)
Unique contacts	-0.01236*** (0.0035)	-0.02809*** (0.0036)

\*\*\*  $p < .001$ .

\*  $p < .05$ .

†  $p < .10$ .

include measures for both undirected and directed contact networks. For the latter,  $i$ 's report of contact with  $j$  may be treated strictly as evidence of  $i$ 's structural orientation to  $j$ , in contrast to the present treatment in which  $i$ 's report of contact with  $j$  suffices to establish a symmetric relation between  $i$  and  $j$ .

The Jaccard-Spline index provides three defensible threshold values,  $p_{ij} > 0.75$ ,  $p_{ij} > 0.50$ , and  $p_{ij} \geq 0.50$ , for defining clusters in  $G$ . With the employment of hierarchical cluster analysis based on the complete-link agglomerative criterion, the first two thresholds will generate complete (not necessarily maximally complete) cliques. Since the clusters are agglomerative in their algorithmic construction, the seeds of the cliques are dyads with the highest structural proximities in  $G$ . With the more stringent of the two thresholds,  $p_{ij} > 0.75$ , the cliques will be complete subgraphs in which each dyad's number of mutual contacts exceeds the dyad's number of unique contacts. With the less stringent threshold,  $p_{ij} > 0.50$ , the cliques may include dyads with more numerous unique than mutual contacts. With the least stringent threshold  $p_{ij} \geq 0.50$ , non-contact dyads may appear in a cluster; however, the criterion for their appearance is stringent, i.e.,  $u_{ij} = 0$  for all such non-contact dyads.

The present findings suggest that instances of cognitive interpersonal solidarity, such as accorded interpersonal influence and perceived interpersonal agreement, are structurally localized relations. It appears (a) that these cognitive relations are constrained to persons who are no more than two-steps removed in  $G$  and (b) that the probability of these relations is conditioned by elementary features of the structure of the contact network—contact, mutual contacts, and unique contacts. The generality of the two-step constraint warrants probing.

A non-contact dyad with many mutual contacts is a structural anomaly, i.e., a low probability event, and the anomaly becomes more pronounced the greater the number of mutual contacts for the non-contact dyad. Something (perhaps antagonism) is inhibiting the formation a contact relation in a non-contact dyad with many mutual contacts. It appears risky to assume that the non-contact dyads of a structurally cohesive group (with a high density of contacts) are pairs of persons with a marked degree of interpersonal solidarity by virtue of their location in such a group. It is perhaps not an oxymoron to acknowledge that a structurally cohesive group may contain interpersonal conflicts that inhibit regular contact between some members. Such structurally anomalous non-

contact dyads are as interesting as local bridges, where a contact occurs between two persons with *no* mutual contacts.

In closing, a comment is warranted on the implications of the present findings for studies of flows of interpersonal influence and the formation of interpersonal agreements. While instances of direct interpersonal influence may be structurally localized in the contact network of a group, indirect interpersonal influences may arise from sequences of these direct influences. Indirect interpersonal influences need not be cognitively acknowledged and manifested in relations of accorded interpersonal influence. Similarly, interpersonal agreements on issues may be formed, via an interpersonal influence process unfolding in a group, which are not manifested in relations of perceived interpersonal agreement. Formal models of social process in networks may draw on structural constructs, e.g., in their specification who is directly responding to whom, but also have implications that extend well beyond the scope of the present article.

## References

- Alba, R.D., 1973. A graph-theoretic definition of a sociometric clique. *Journal of Mathematical Sociology* 3, 113–126.
- Ahlberg, J.H., Nulson, E.N., Walsh, J.L., 1967. *The Theory of Splines and their Applications*. Academic, New York.
- Batagelj, V., Bren, M., 1995. Comparing resemblance measures. *Journal of Classification* 12, 73–90.
- Brass, D.J., 1984. Being in the right place: a structural analysis of individual influence in an organization. *Administrative Science Quarterly* 29, 518–539.
- Burt, R.S., 1978. Cohesion versus structural equivalence as a basis for network subgroups. *Sociological Methods & Research* 12, 189–212.
- Burt, R.S., 1987. Social contagion and innovation: cohesion versus structural equivalence. *American Journal of Sociology* 92, 1287–1335.
- Burt, R.S., 1999. The social capital of opinion leaders. *The Annals of the American Academy of Political and Social Science* 566, 37–54.
- Erickson, B.H., 1988. The relational basis of attitudes. In: Wellman, B., Berkowitz, S.D. (Eds.), *Social Structures: A Network Approach*. Cambridge University Press, Cambridge, pp. 99–121.
- Feld, S.L., 1981. The focused organization of social ties. *American Journal of Sociology* 86, 1015–1035.
- Freeman, L.C., 1979. Centrality in social networks: conceptual clarification. *Social Networks* 1, 215–239.
- Friedkin, N.E., 1984. Structural cohesion and equivalence explanations of social homogeneity. *Sociological Methods & Research* 12, 235–261.
- Friedkin, N.E., 1986. Horizons of observability and limits of informal control in organizations. *Social Forces* 62, 54–77.
- Friedkin, N.E., 1998. *A Structural Theory of Social Influence*. Cambridge University Press, Cambridge.
- Friedkin, N.E., 2004. Social cohesion. *Annual Review of Sociology* 30, 409–425.
- Granovetter, M.S., 1973. The strength of weak ties. *American Journal of Sociology* 78, 1360–1380.
- Luce, R.D., Perry, A., 1949. A method of matrix analysis of group structure. *Psychometrika* 14, 95–116.
- McPherson, M., Smith-Lovin, L., Cook, J., 2001. Birds of a feather: homophily in social networks. *Annual Review of Sociology* 27, 415–444.
- Meyer, G.W., 1994. Social information processing and social networks: a test of social influence mechanisms. *Human Relations* 47, 1013–1047.
- Mizruchi, M.S., 1993. Cohesion, equivalence, and similarity of behavior: a theoretical and empirical assessment. *Social Networks* 15, 275–307.
- Mokken, J.C., 1979. Cliques, clubs and clans. *Quality and Quantity* 13, 161–173.
- Padgett, J.F., Ansell, C.K., 1993. Robust action and the rise of the medici, 1400–1434. *American Journal of Sociology* 98, 1259–1319.
- Roethlisberger, F.J., Dickson, W.J., 1939. *Management and the Worker*. Harvard University Press, Cambridge, MA.
- Seidman, S.B., Foster, B.L., 1978. A graph-theoretic generalization of the clique concept. *Journal of Mathematical Sociology* 6, 139–154.
- White, H.C., Boorman, S.A., Breiger, R.L., 1976. Social structure from multiple networks I: blockmodels of roles and positions. *American Journal of Sociology* 81, 730–781.