

E-Companion to
WHAT DO THEY KNOW? THE ANTECEDENTS OF INFORMATION ACCURACY
DIFFERENTIALS IN INTERORGANIZATIONAL NETWORKS

By **J. Knoben, L.A.G. Oerlemans, A.R. Krijkamp & K. Provan**

Calculation of Network Information Accuracy at the Directed Pair Level

This Appendix provides a detailed explanation of the calculation of the network information accuracy variable at the level of directed pairs. There is a myriad of measures that can be used to assess the degree to which an organization's perceptions of links between other organizations links correspond to the actual existence of those links (Krackhardt, 1990). All these measures rely on the information depicted in the following 2x2 table.

		<i>Objective relations of organization 'j'</i>	
		<i>Absent</i>	<i>Present</i>
<i>Perception of organization 'k' of relations of organization 'j'</i>	<i>Absent</i>	<i>a</i>	<i>b</i>
	<i>Present</i>	<i>c</i>	<i>d</i>

Following a review by Gower and Legendre (1986) of 15 possible measures that capture the correspondence in such 2x2 tables Krackhardt (1990) opted to use a measure called 's14'. This measure was chosen for its high resolution (appropriate sensitivity to small changes in correspondence) and low nonlinearity (low distortion at extreme values). Mathematically 's14' takes the following form:

$$s14 = \frac{ad - bc}{\sqrt{(a + c)(b + d)(a + b)(c + d)}}$$

Ideally, we would have followed Krackhardt's measurement of accuracy. However, 's14' is problematic to calculate at the directed pair level. Specific for our case it cannot be calculated when 'a + b = 0' or 'a + c = 0'. When assessing perceptions of an entire network, as in Krackhardt (1990), such situations are rare and only occur either when there are no ties present in the network (a + c = 0) which means that a network does not exist, or when an organizational respondent perceives the network as containing zero ties (a + b = 0), which would reflect an extreme case of lack of knowledge. At the directed pair level which we use, however, such situations can readily occur, as when an assessed organization is an isolate (a + c = 0) or when the organization is perceived to be an isolate (a + b = 0). Our networks do contain some isolates and several organizations are perceived to be isolates. Thus, s14 is not calculable at the directed pair level.

We relied on the same review by Gower and Legendre (1986) to select a measure with high resolution and low non-linearity that could be calculated at the directed pair level.¹ Ultimately, we opted for the measure 's9' from the Gower and Legendre (1986) paper. The measure we adopted has an even higher resolution than the one used by Krackhardt (1990) and only a marginally higher non-linearity. Formally, our measure is expressed as:

¹ Another possibility would be to apply an inter-coder reliability measure such as Cohen's Kappa or Scott's Pi. However, these types of measures have been shown to be extremely sensitive to small changes in cases where the observations are highly unevenly spread over the four possible cells (de Mast, 2007). This is almost by definition the case in network research where most organizations have few relations and some organizations have many. As such, we opted not to use these measures.

$$s_9 = \frac{a + d - (c + b)}{a + b + c + d}$$

To assess how our measure compares to the one used by Krackhardt (1990) we calculated both at the organizational level. The correlation between the two measures was 0.83. Since both measures have similar characteristics, similar range and are highly correlated at the higher level of analysis, we conclude that this measure is the best measures of accuracy that can be constructed at the directed pair level.

To assess the sensitivity of our results to the specification of the dependent variable we ran our models with several different measures of accuracy. Specifically, we used the measures s4 and s9 from the Gower and Legendre (1986) paper as well as a custom combination of these measures called s4/9. For each of these we find highly similar results (see Table A1 below) indicating that our findings are rather robust to changes in the specification of the dependent variable.

Table A1

Regression Analyses For Various Network Information Accuracy Measures^a

Variables	S9 = (a+d-b-c) / (a+b+c+d)		S4 = (a+d) / (a+b+c+d)		S4/9 = (a+d-b) / (a+b+c+d)		S6 = (a+d) / (a+2(b+c)+d)	
Size of the focal organization	-0.01	0.01	0.01	0.01	0.01**	0.00	0.01	0.01
Subsidiary status focal organization	-0.03**	0.01	0.01	0.02	0.03*	0.02	0.01	0.01
Religious mission focal organization	-0.02	0.02	-0.02†	0.01	-0.06**	0.02	-0.03**	0.01
Size of the assessed organization	-0.01	0.01	-0.01	0.01	-0.01	0.01	-0.01	0.01
Subsidiary status assessed organization	-0.02	0.02	-0.02	0.02	-0.02	0.02	-0.01	0.01
Religious mission assessed organization	-0.04†	0.02	-0.02†	0.01	-0.02	0.02	-0.02†	0.01
Direct tie	-0.03***	0.01	-0.02**	0.01	-0.03**	0.01	-0.03***	0.01
Geographical proximity	0.01†	0.01	0.01*	0.00	0.01*	0.00	0.01	0.01
Domain similarity	0.30	1.12	0.33	0.68	0.86	0.94	0.22	0.69
Degree of the focal organization	0.05***	0.01	0.05***	0.01	0.08***	0.01	0.03***	0.01
Degree of the focal organization squared	-0.02**	0.01	-0.03	0.03	-0.01***	0.00	-0.01**	0.00
Degree of the focal organization * Domain similarity	-0.03	0.06	0.25	0.56	0.68	0.82	0.15	0.60
Degree of the focal organization squared * Domain similarity	0.03**	0.01	0.01**	0.00	0.03*	0.01	0.03**	0.01
Degree of the assessed organization	-0.03***	0.01	-0.02***	0.00	-0.03***	0.01	-0.03***	0.00
Degree of the assessed organization * Domain similarity	-0.01	0.20	-0.05	0.11	-0.14	0.15	-0.04	0.11
Type of activities fixed effects	Yes		Yes		Yes		Yes	
Network level fixed effect	Yes		Yes		Yes		Yes	
Number of observations	2052		2052		2052		2052	
Model significance	0.00		0.00		0.00		0.00	
R-squared	38.21%		26.69%		33.60%		46.97%	

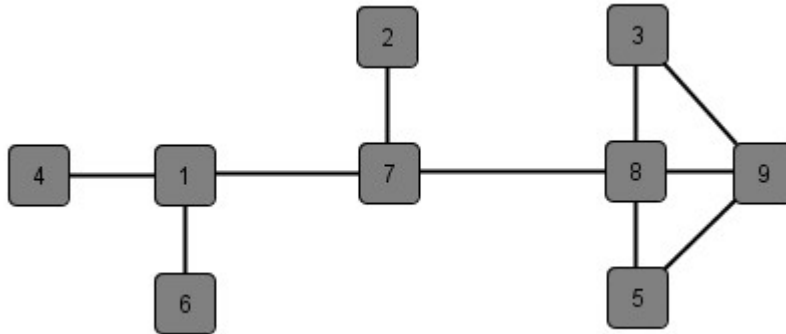
a: Robust standard errors clustered at the focal organizational and the assessed organizational level are in parentheses

† p < .10; * p < .05; ** p < .01; *** p < .001

We proceed by giving an example calculation of our accuracy variable. For reasons of clarity we do these calculations using a small, hypothetical, network, and assume that the objective network structure is as depicted

in Figure A1. Within this network, we examine accuracy of the perception of node 1 (focal node) regarding the network position of node 8 (assessed node). To emphasize that the network information variable is directed, we will also calculate the perception accuracy score of node 8 (focal node) regarding the network position of node 1 (assessed node).

Figure A1. Objective Network Structure



The perception of the network by node 1 is depicted in Figure A2. Since we are calculating the accuracy of the perception of node 8's network position, we are only interested in the bold ties. A comparison of figure A1 and A2 reveals that on the one hand, node 1 correctly recognizes that node 8 has ties with nodes 3, 5, 7, and 9. However, node 1 also incorrectly perceives node 8 to be connected to nodes 2 and 6. Finally, node 1 correctly perceives no ties between node 8 and nodes 1 (itself) and 4. This information is summarized in Table A2. Calculating the accuracy score following the formula given in the paper yields: $Accuracy_{1_8} = \frac{4+2-(0+2)}{4+2+2} = 0.50$.

Figure A2. Perception of Network Structure of Node 1

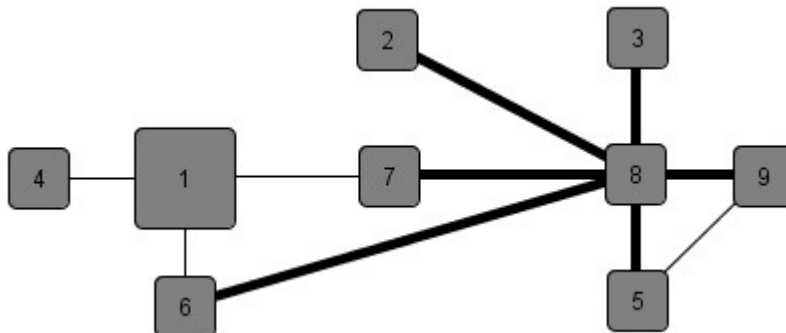


Table A2. Classification Table of Node 1's Perception of Node 8's Network Position

		Objective relations of organization 8	
		Absent	Present
Perception of node 1 of relations of organization 8	Absent	2	0
	Present	2	4

It is important to note that the network information accuracy of node 8 regarding node 1's network position is independently calculated. Figure A3 depicts the network perception of node 8. Again, we are only interested in the bold ties since we are assessing the accuracy of node 8's perception of node 1's network position. In this case, node 8 (correctly) perceives node 1 to be connected to node 7 but fails to report the connections between node 1 and nodes 4 and 6. This yields a classification of ties as reported in Table A3 and an accuracy score of: $Accuracy_{8_1} = \frac{5+1-(2+0)}{5+2+1} = 0.5$

Figure A3. Perception of Network Structure of Node 8

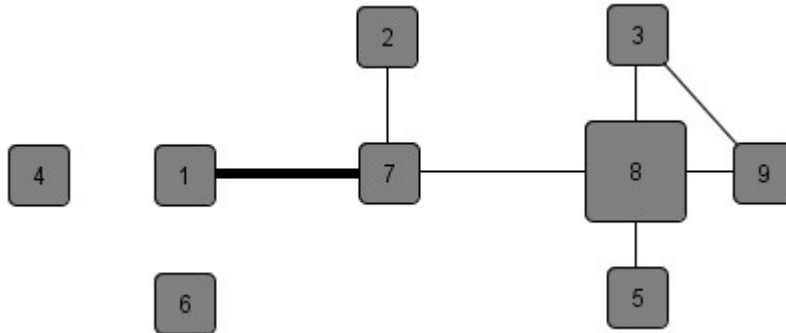


Table A3. Classification Table of Node 8's Perception of Node 1's Network Position

		<i>Objective relations of organization 1</i>	
		<i>Absent</i>	<i>Present</i>
<i>Perception of node 8 of relations of organization 1</i>	<i>Absent</i>	5	2
	<i>Present</i>	0	1

In our dataset, these two scores represent the dependent variable of two separate observations. The structure of the dataset that follows from the above is depicted in Table A4. However, it should be noted that even for our small hypothetical network of 9 nodes used in this Appendix, the full dataset would contain 72 directed pairs.

Table A3. Illustration of Dataset Structure

Focal node	Assessed node	Accuracy	Degree focal	Degree Assessed
1	8	0.50	3	4
8	1	0.50	4	3

The observations in the dataset are in some respects each other's mirror image. For example, the degree centrality of the focal organization for the first observation is identical to the score of the degree centrality of the assessed organization for the second observation, and vice versa. The dependencies between the observations that arise as a result of this data structure are the reason we need to utilize the multiway clustering approach (Cameron et al., 2011) when analyzing the data, as described in the 'Statistical Analyses' section of the paper.