How many is good enough? A study for the nomination of a social network

Hsieh-Hua Yang
Chang Jung Christian University

Chyi-in Wu
Man Kit Lei
Institute of Sociology, Academia Sinica

Paper presented at the “調查研究方法與應用”, Sep 2-4, 2004, Taipei, Taiwan.

Abstract

For social network analysis, the most common method used to generate social network data is the method of Name Generator. Researchers usually ask the respondents, generally affiliated with each other within an organization, a class or a group with clearly boundary, to name several their best friends overall or people they mostly contacted during a certain period in a spread sheet. Later, the researchers can translate this information into a person-by-person data matrix to do social network analysis. Nonetheless, the number of nomination has been an unsolved mystery for social networkers. Namely, for each respondents, to name how many people is good enough to generate a stable network, which is able to represent the truly association structure among these respondents, still, is an empirical research question for researchers.

This study devoted to explore this question and to provide a preliminary answer. A set of social network data was collected from a sample of Taipei metropolitan schools, including 44 classes. In each class the students were asked to nominate ten best friends in the intimate order. It was supposed that in each class has ten sociometric data for different nomination, and the total amount of sociometric data was 440. KrackPlot3.20 was applied to layout the social network graph, UCINET5.0 to analyze the social network variables, and NEGOPY4.30 to define the network position. Visualizing the graph of each class and comparing the betweenness, constraint, and efficiency, this study found that two names will generate more diverse network position with unstable structure, three names are the minimum to get more stable network structure, four or five names are needed to observe the links between boys and girls, but more than five names seem to be redundancy.

1. Introduction

Adolescents are embedded in a rich network of peer relations (Furman, 1989), including best friendships, cliques or friendship groups. Studies of adolescent peer relations have focused on the characteristics of peer social network, and its correlations with deviant behavior. The most common method used to gather network data is name generator. Subjects are asked to name their best friends. The number of friends named either has a limitation of three to ten or no limitation at all.

Many researchers asked the subjects to name three friends in studying friendship structures. Fisher and Bauman (1988), Ennett and Bauman (1994) focused on the influence and selection in the adolescent peer group. Ennett and Bauman (1996) studied the variability of friendship structures. While Urberg, Degirmencioglu, Tolson, and Halliday-Scher (1995) provided 10 spaces for listing friends to study the structure of adolescent peer networks, but only the mutual choices were included in analysis. Cairns, Leung, Buchanan, and Cairns (1995) studied the fluidity, reliability, and interrelations of children's and adolescents' social networks. Subjects were free to nominate any number of friends.

Others analyzed the adolescent social network to establish an association between peer group structure and deviant behavior. Bauman and Fisher (1986), Ennett and Bauman (1993), Aloise-Young, Graham, and Hansen (1994) studied the peer influence on smoking behavior; they asked the adolescents to name three best friends. Yang, Chen, Yen, and Ke (2002, 2003) asked high school students to name four friends in the same class to explore the effects of peer group on the smoking behavior. Michell and Amos (1997) asked pupils to name up to six friends to study the girls' smoking behavior. Ennett, Bailey & Federman (1999) studied the social network characteristics associated with risky behaviors among runaway and homeless youth. The youths provided information on up to seven members of their networks. Urberg, Degirmencioglu, and Pilgrim (1997) studied the close friend and group influence on adolescent cigarette smoking and alcohol use. Adolescents were asked to nominate up to 10 best friends. Haynie's (2001) studied the impact of peer influence on adolescent delinquency. Adolescents were asked to identify their best male and best female friends from a school roster (up to 5 friends of each sex).

Most of these studies didn't discuss how to decide the number of nomination. Pearson and Michell (2000) explained why they asked the subjects to give information on up to six friends. Their motivation to use six friends was the need to strike a balance between naming too few links which generates peer structures which are inadequately cohesive and having too many choices which tends to provoke subjects to name peers who are not close friends leading to an overestimation of the number of links.

Our focus on the number of nomination is appropriate because no prior study has empirically examined this issue. In this study an empirical data was analyzed to examine individual network variables and network structure for finding the best number of nomination.

2. Method

A random sample of 1,434 junior high school students was collected from 44 classes of 33 schools at Taipei. The 33 schools were selected randomly to provide a socially and economically diverse sample of the whole Taipei metropolitan city. And in these schools, 44 classes were randomly selected from the 7th grade during 1996.

The students were asked to nominate 10 best friends in terms of intimacy. The boundary of network was limited in the same class. A nominated friend not in the same class was not included in the analysis. We applied KrackPlot 3.3 (Krackhardt, Blythe, and McGrath, 1994) to layout the social network graph, UCINET 5.0 (Borgatti, Everett, and Freeman, 1999) to analyze individual network variables, and NEGOPY4.30 (Richards, 1995) to define the position in the network. There are a series of social network graphs from one to ten nominations in each class. And for each social network, there are individual network variables including betweenness[1], constraint[2], and efficiency[3], and network structure variables including a variety of network position.

Many writers have offered definition of degree, closeness, betweenness, and density as individual network variables. Degree which is defined as the total number of direct connections and closeness as a point's geodesic distances to all points on the graph (Scott, 2000; Degenne and Forse, 1999; Wasserman & Faust, 1994) are relative to nomination number. Density is a size-dependent measure which is difficult to use in comparisons of graphs of radically different sizes (Scott, 2000). While betweenness measures the extent to which a particular point lies 'between' the various other points in the graph and a point of relatively low degree may play an important 'intermediary' role (Scott, 2000). It is not relative to nomination number and may be used as a good measure. And for comparing between different classes, we normalized the betweenness as our measure. Constraint and efficiency were offered by Burt (1992). Constraint is essentially a measure of the extent to which ego is invested in people who are invested in other of ego's alters. Efficiency is the effective size divided by the number of alters in ego's network. They are not relative to nomination number directly and were computed as a measure.

KrackPlot 3.20 is a social network analysis program which offers a pictorial image of social network. The nodes can be specified with different style to present its characteristics. Here the box was use to present girls in a network and the different style would help us to observe the interaction between boys and girls.

---

[1] ***CB(ni)= Σbjk(ni)/bjk*** Let bjk be the proportion of all geodesics linking vertex j and vertex k which pass through vertex i. The betweenness of vertex i is the sum of all bjk where i, j and k are distinct and j < k. Betweenness is therefore a measure of the number of times a vertex occurs on a geodesic. The normalised betweenness centrality is the betweenness divided by the maximum possible betweenness expressed as a percentage.

[2] ***Constraint*** $= (p_{ij} + \sum p_{iq}p_{qj})^2$ p_{ij}=proportion of i's relations allocated to j. p_{iq}=proportion of i's energy invested in relationship with q. p_{qj}=i's interaction with q divided by j's strongest relationship with anyone.

[3] ***Efficiency = Effective size/observed size, Effective size*** $= \sum [1-\sum p_{iq}m_{jq}]$ m_{jq}=i's interaction with q divided by j's strongest relationship with anyone. p_{iq}=proportion of i's energy invested in relationship with q. ***Effective size*** is network size (N) minus redundancy in network.

UCINET 5.0 is software of social network analysis created by Borgatti, Everett, and Freeman. We applied it to analysis individual network variables, including betweenness, constraint, and efficiency.

NEGOPY is another social network analysis program. It detects groups and assigns individuals in the network to a discrete set of categories. This set of categories is based on a systems-theoretic approach to organizations. There are two major categories into which individuals are assigned—isolates and participants. Isolates include all the individuals who are minimally connected to others in the network and have four types. Isolate type 1 are the individuals who have no links to any other individuals in the network. Isolate type 2 are all the individuals who are linked to only one other individual in the network. If two isolates type 2 are linked to one another, they are called a dyad. And if a set of individuals are connected in a structure that is acyclic (contains no cycles), they are tree nodes.

Participants are individuals who have at least two links with other participants. They include both members of groups and individuals who connect groups to one another. Group members have most of their interaction with other members of the same groups, rather than with members of other groups. Direct liaisons are individuals who have most of their interaction with members of groups, but not with members of any one group. They provide direct connections between the groups they are connected to. And indirect liaisons are individuals who do not have most of their interaction with members of groups. They provide indirect of 'multi-step' connections between groups by connecting liaisons, which have direct connections with members of groups.

The difference between isolates and participants is the number and the form of linkage with others in a network. When the number of nomination increased, an isolate type 1 become type 2 or dyad, an isolate type2 become a liaison or a group member, dyad become tree nodes, tree nodes become group member, et cetera. It implies that the number of nomination will change the network structure. And we supposed that the increased nomination would not be necessary if it had not made the network structure any change.

For the purpose of observing the change of network positions, a series of social networks were analyzed. Now, in each class there were 10 networks from 1 to 10 nominations. We observed the categories of social network positions, compared the difference between two adjacent social networks, and calculated the change rate of each category of network positions. If there were no difference between two adjacent social networks, we supposed the latter social network (more number of nominations) was structural stable, and the number of former social network is the good number of nomination.

3. Findings

There were 727 boys (50.7%) and 707 girls (49.3%) in 44 classes, including 6 boys' classes, 6 girls' classes, and 32 boys and girls' classes. The results of analyze are summarized in three sections. In the first part, the individual network variables are described. The second part focuses on network structure. And the third part, the individual network variable and network structure are linked.

3-1. Individual network variables

The frequency of nomination number was shown as table 1. The mean of nominations was 4.30 (SD=2.94). Among 1434 subjects, 109 (7.6%) students named up to 10 friends, and 138 (9.6%) didn't name any friend. Most of the subjects named one to five friends, 145 (10.1%) students named 1 friends, 162 (11.3%) students named 2 friends, 206 (14.4%) students named 3 friends, 163 (11.4%) named 4 friends, 158 (11.0%) named 5 friends.

The variances of betweenness were counted for each different nomination. For each class there was a variance curve from 1 to 10 nominations. As shown in figure 1, in 44 classes there were two classes which curves were higher than others. The individual betweenness in these two classes were shown as figure 1-1 and 1-2. From these figures we could find whose betweenness changed sharply. They were "687" and "695" in class A, "989" and "978" in class B.

The variances of constraint and efficiency of 44 classes were consistently decreased as the nomination was increased from 1 to 3. The curves became flat as the nomination number was increased to 4 or 5. (Figure 2 and 3)

3-2. network structure

We used NEGOPY to defined network position for each student within their classes. There are two major categories, participants and isolates. Participants include group members and two kinds of liaisons. Isolates include two kinds of isolates, dyad, and tree nodes. The network position will be changed while the nomination number increased. We calculated the number of network position for different nomination, and compared the difference between two different nominations. The result is shown as table 2.

The change rate of group members from 2 to 3 nomination was 25.77%, it meant that there are about a quarter of subjects became group members as the nomination increased from 2 to 3. While the group members increased, the isolates decreased. The decreased rates were 5.11%, 5.84%, 2.63%, and 9.58% for isolate type 1, isolate type 2, dyad and tree nodes respectively. At the same time, the direct liaisons were decreased by 2.64%.

From 3 to 4 nomination, the change rate of group members was 9.01%, meanwhile the change rates of isolates were -3.80%, -1.36%, -0.68%, and -2.73% for each isolates category (Table 2). And from 4 to 5, or more nominations, the change rates were less than ±5% for each category of network positions (Figure 4).

The increase of nomination would change the categories of network position. As the nomination number was only one, there were not any participants. Most of the students were isolates. While the nomination number was more than 5, most of the students became group members. Between one and five nominations, we found 2 nominations would generate more diverse network position than any other nominations for most of these classes, as shown in table 3.

Comparing the adjacent network structure, we identified the nomination number of structural stable as presented in table 4. There were three criteria for this comparison. For criteria I, network position was classified as group member, dyad, tree node, isolate type 1, isolate type 2, liaison type 1, and liaison type 2. As shown in the first column of table 4, there were 9 (20.45%) classes got structural stable as the nomination increased to 5. Eight (18.18%) classes got structural stable as the nomination increased to 6. And 6 (13.64%) classes didn't get structural stable as the nomination increased to 10. For criteria II, shown in the second column of table 4, network position was classified as participants and isolates. There were 7 (15.91%) classes got stable as the nomination number was 4, 12 (27.27%) classes got stable as the nomination increased to 5, and 10 (22.73%) classes got stable as increased to 6. For criteria III, shown in the third column of table 4, network position was classified as group members and not group members. Most of the classes' networks got structural stable as the nomination number increased to 6. There are 7 (15.91%) classes got stable as the nomination number was 4, and 17 (38.64%) classes got stable as the nomination number increased to 5, and 10 (22.73%) classes got stable as increased to 6.

3-3. individual network variable and network structure

There were two classes with extraordinary high variances of betweenness. We further examined these two classes' network map. These two higher curves are boys and girls' classes. In these two classes, boys and girls were separated into different components until the nomination number increased to 4 or 5. A boy nominated a girl or a girl nominated a boy would become a bridge between boys and girls and increase the betweenness. In class A, as figure 5, when the nomination number increased to 4, 695 (a girl) and 687 (a boy) linked, and their betweenness increased (Figure 1-1). In class B, as figure 6, a boy (989) only had a dyad link with a girl (978). When the nomination number increased to 5, the boy's and the girl's outdegree and indegree increased. Their betweenness increased as shown in figure 1-1.

4. Discussion

The mean of nomination was 4.30. Most of them named one to five friends. Only 7.6% named 10 friends. The result was very close to others who asked the subjects to name 10 friends. The mean number of nomination in Haynie's research was 4.15 (SD=3.02). And the average number of friends listed in Urberg, Degirmencioglu, and Pilgrim's studies was 4 to 5. In other studies which asked subjects to name three friends, the mean size was 2.6 (Ennett, Bailey & Federman) and 41% of adolescent named three friends in Ennett's study (1996). It looks like that the more asked the more named. But when the number increased to more than the ultimate number of subjects' friends, the number of nomination will not be increased any more. In Cairns, Leung, Buchanan, and Cairns's study, subjects were free to nominate any number of friends. The result showed that a mean of 4.09 friends were named in the seventh grade. Comparing with these findings, we suggest that 4 or 5 friends may be an acceptable number for junior subjects to nominate.

Betweenness measures the extent to which a particular point lies 'between' the various other points in the graph (Scott, 2000). We scrutinized the network map with higher variances of betweenness, and found that the junior students did keep friends with the same gender. The linkage between boys and girls began at the nomination number increased to 4 or 5. It tells us that if we want to observe the interaction between boys and girls within a class, the nomination number should be up to 4 or 5.

The network position diversified as the nomination number was 2, the network structure changed dramatically as the nomination number increased from 2 to 3, and kept steady when the nomination number increased to 4 or 5. This finding suggested that if we want to get more diversity, 2 names are enough. But the network structure is unstable, and may not represent the truly association structure. If we want to have a more stable structure, 3 names are the lower limit. Additionally, the variances of constraint and efficiency displayed the same characteristics. The variances were changed during nomination number increased from 1 to 3, and kept steady as the nomination number increased to 4 or 5. It also supported that more than 4 or 5 nomination would generate more redundant links, and let the variation of constraint and efficiency for each subject vanished.

Network graph will tell us the same thing more clearly. Look at figure 5, the network map of class A, when the students named two friends, there are 3 isolates type 1, 6 isolates type 2, 2 dyad members, 5 tree nodes, and 3 group members. When the students named three friends, the original 2 tree nodes and 2 isolates type 2 formed a group, another tree node, 684, entered into a group. And the number of group members increased. But when the named friends increased to five, there had little change occurred. One of the isolates type 2 entered a group, and one of the tree nodes became a direct liaison. As the network structure keeps steady, the increased nominations will not be necessary to detect the network positions. The phenomenon suggested that three names seem to be a minimum because it needed to reveal variation in network structure, and more than five nominations seem to be redundancy.

Table 1. Frequency of outdegree

Outdegree	N	%
0	138	9.6
1	145	10.1
2	162	11.3
3	206	14.4
4	163	11.4
5	158	11.0
6	116	8.1
7	112	7.8
8	61	4.3
9	64	4.5
10	109	7.6
Total	1434	100.0

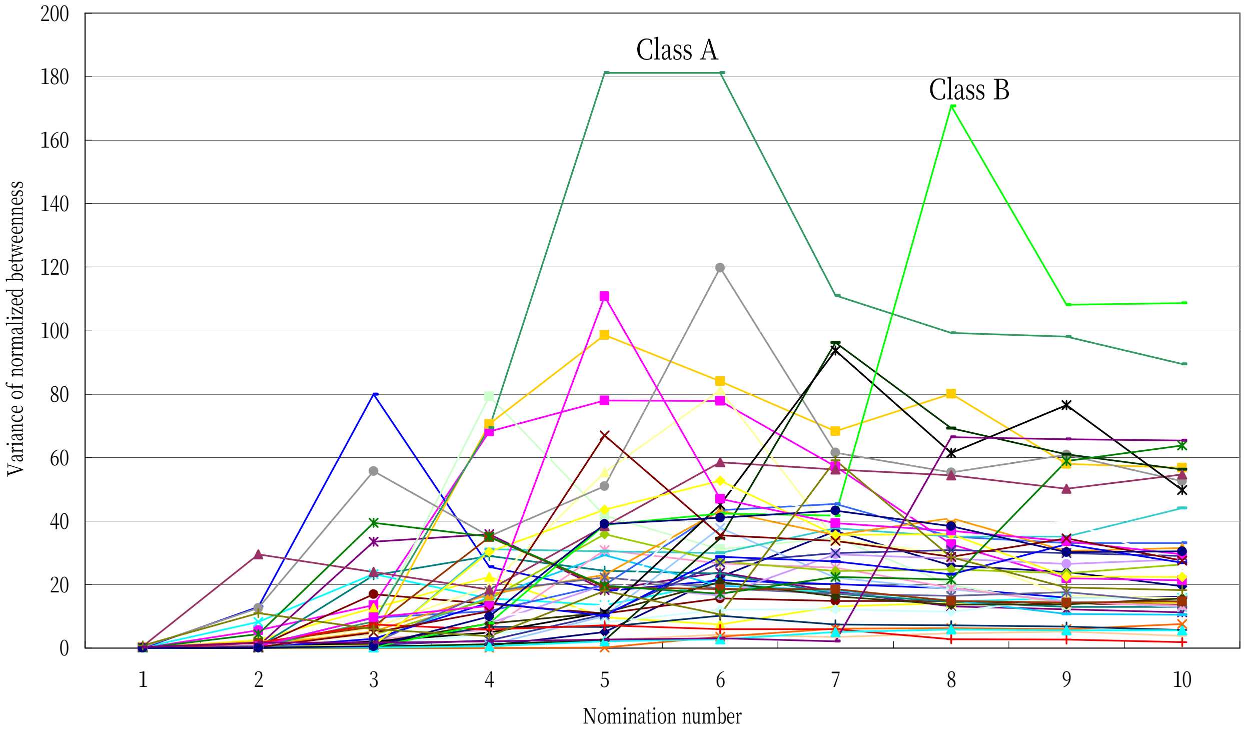
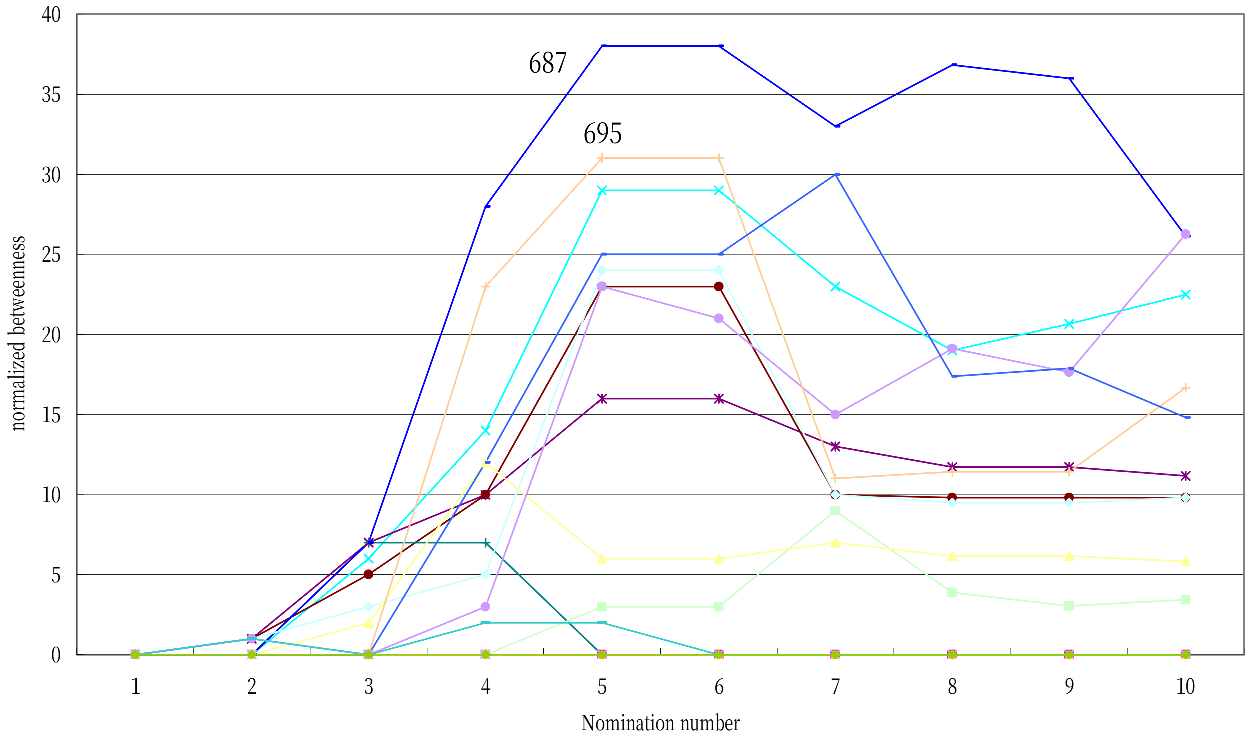
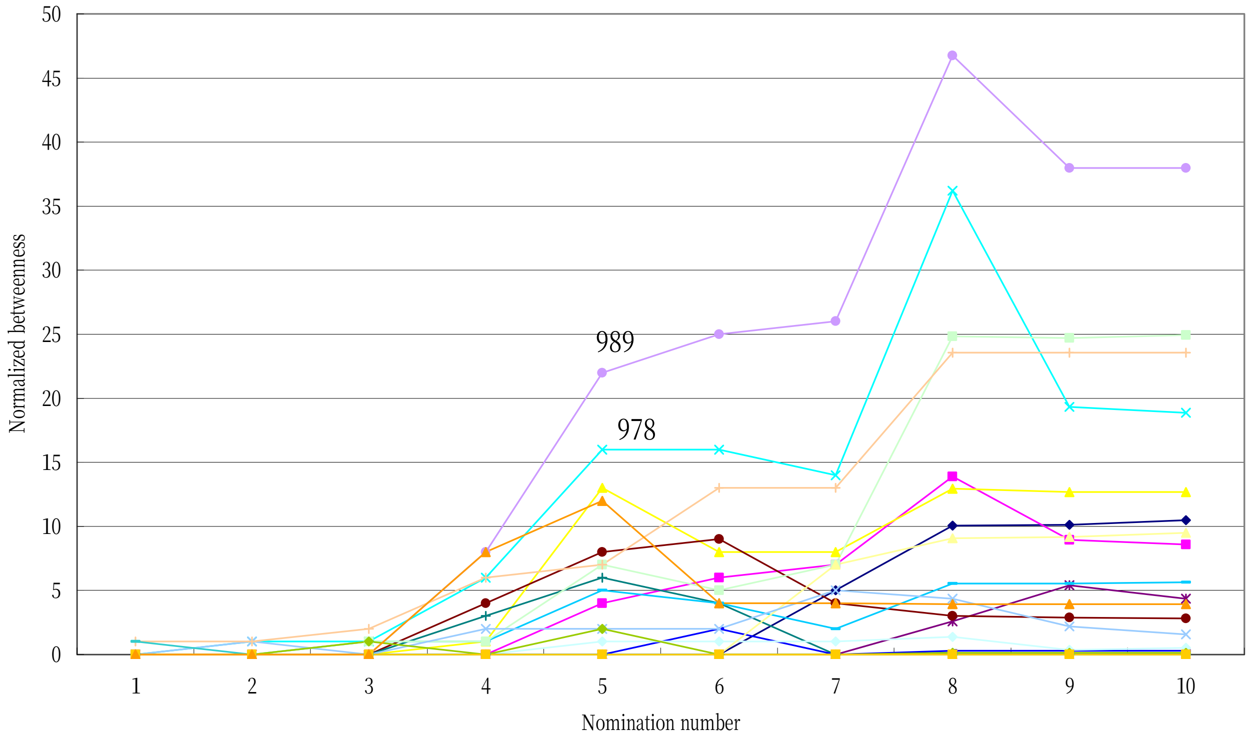


Figure 1. Variance of betweenness of 44 classes



Note: 687 is a boy, and 695 is a girl in class A.

Figure 1-1. Normalized betweenness in class A (12 boys and 7 girls)



Note: 989 is a boy, and 978 is a girl in class B.

Figure 1-2. Normalized betweenness in class B (8 boys and 13 girls)

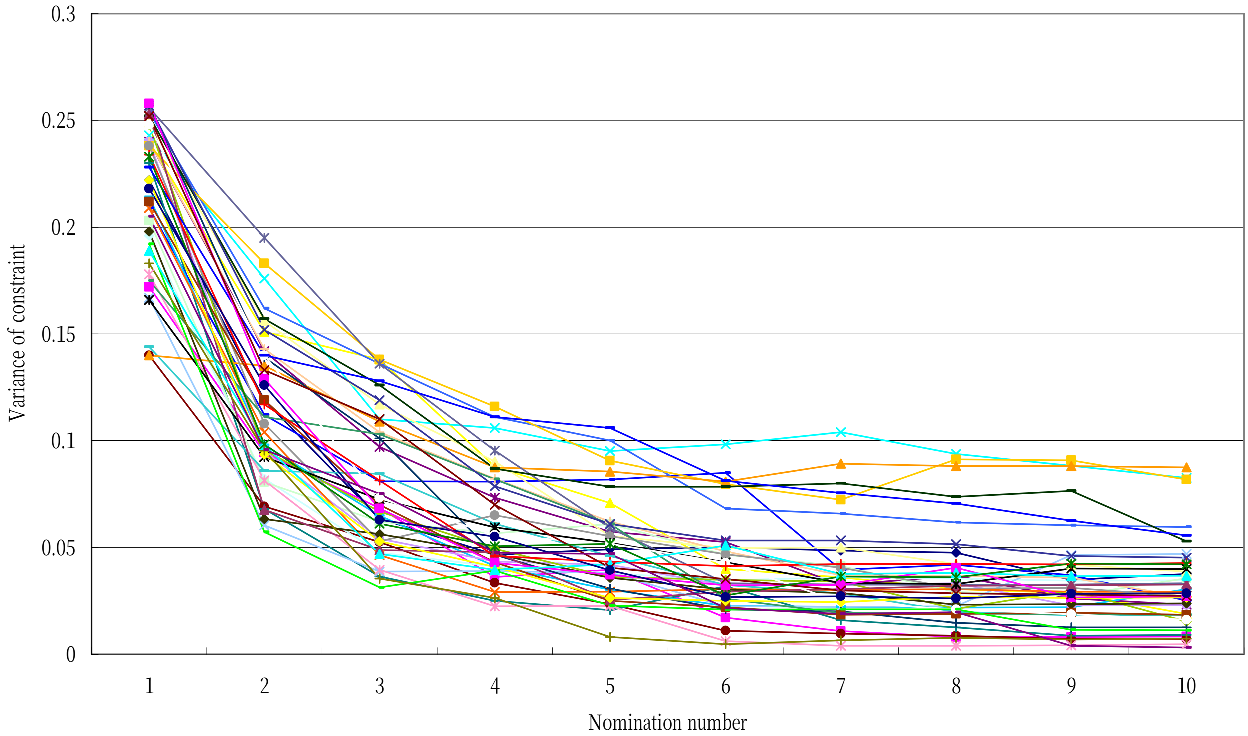


Figure 2. Variance of constraint of 44 classes

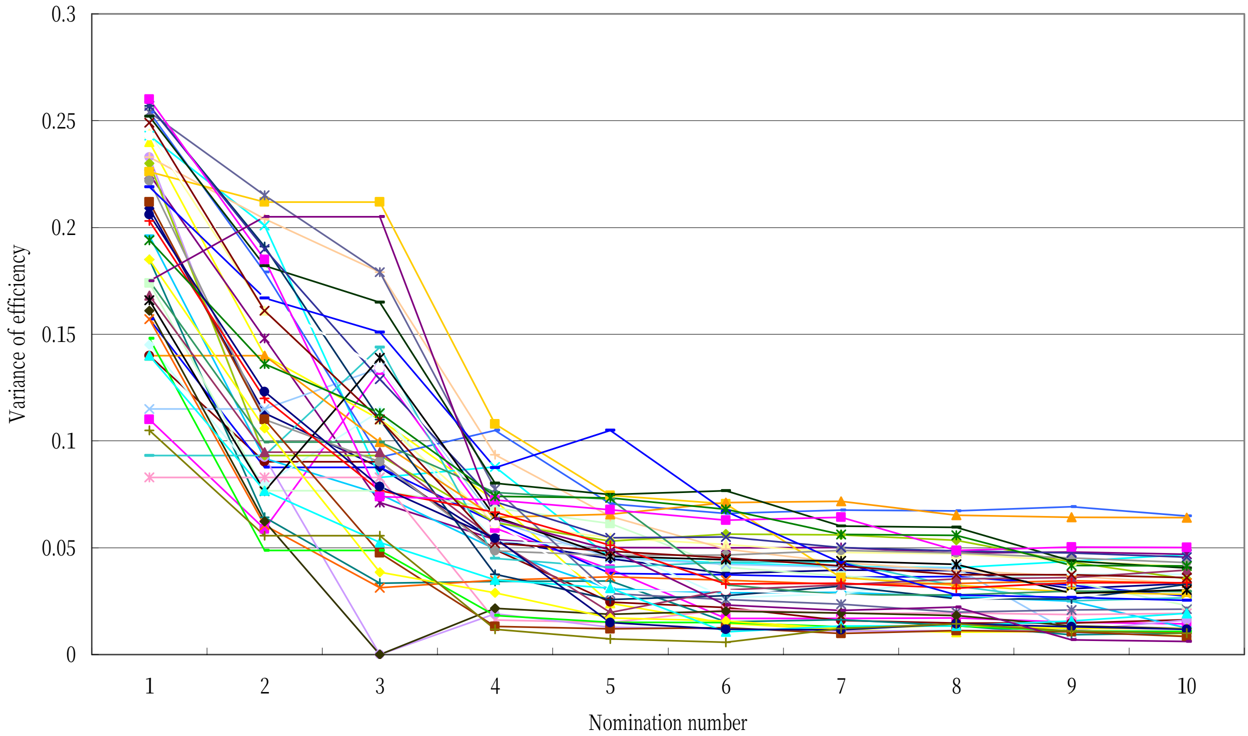


Figure 3. Variance of efficiency of 44 classes

Table 2. Change rate of network position between different nominations

Nomination number	Group member	Isolates (type1)	Isolates (type2)	Dyad	Tree node	Liaisons (direct)	Liaisons (indirect)
2→3	25.77	-5.11	-5.84	-2.63	-9.58	-2.64	.02
3→4	9.01	-3.80	-1.36	-.68	-2.73	-1.13	.71
4→5	4.49	-2.07	-2.11	-.04	-1.31	.99	-.17
5→6	4.05	-.38	-2.58	-.13	-.16	-.94	.37
6→7	1.66	-.45	-.29	-.28	.02	.23	-.89
7→8	1.27	-.15	-.78	.00	-.15	-.19	-.13
8→9	.98	-.37	-.06	-.02	.07	-.85	.25
9→10	-.18	-.11	-.61	.00	-.13	-.09	1.11

Note: 2→3 means the nomination number increased from 2 to 3.

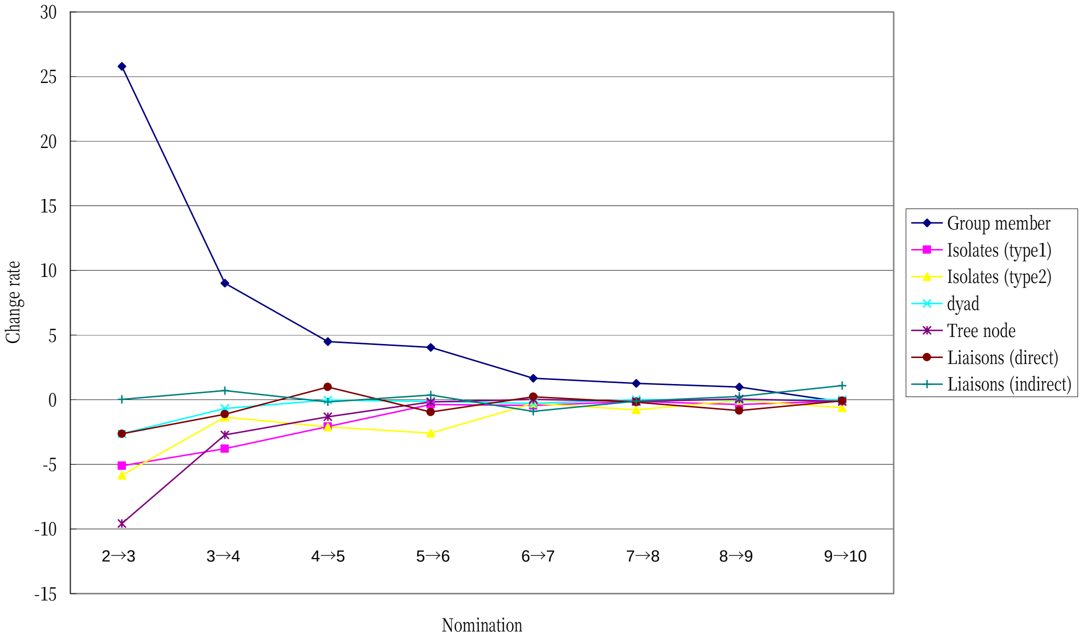


Figure 4. Change rate of network position

Table 3. Nomination number for diversity of network position categories

Nomination number	No. of classes	%
2	33	75.00
3	6	13.64
4	3	6.82
6	2	4.55

Table 4. Nomination number of structural stable

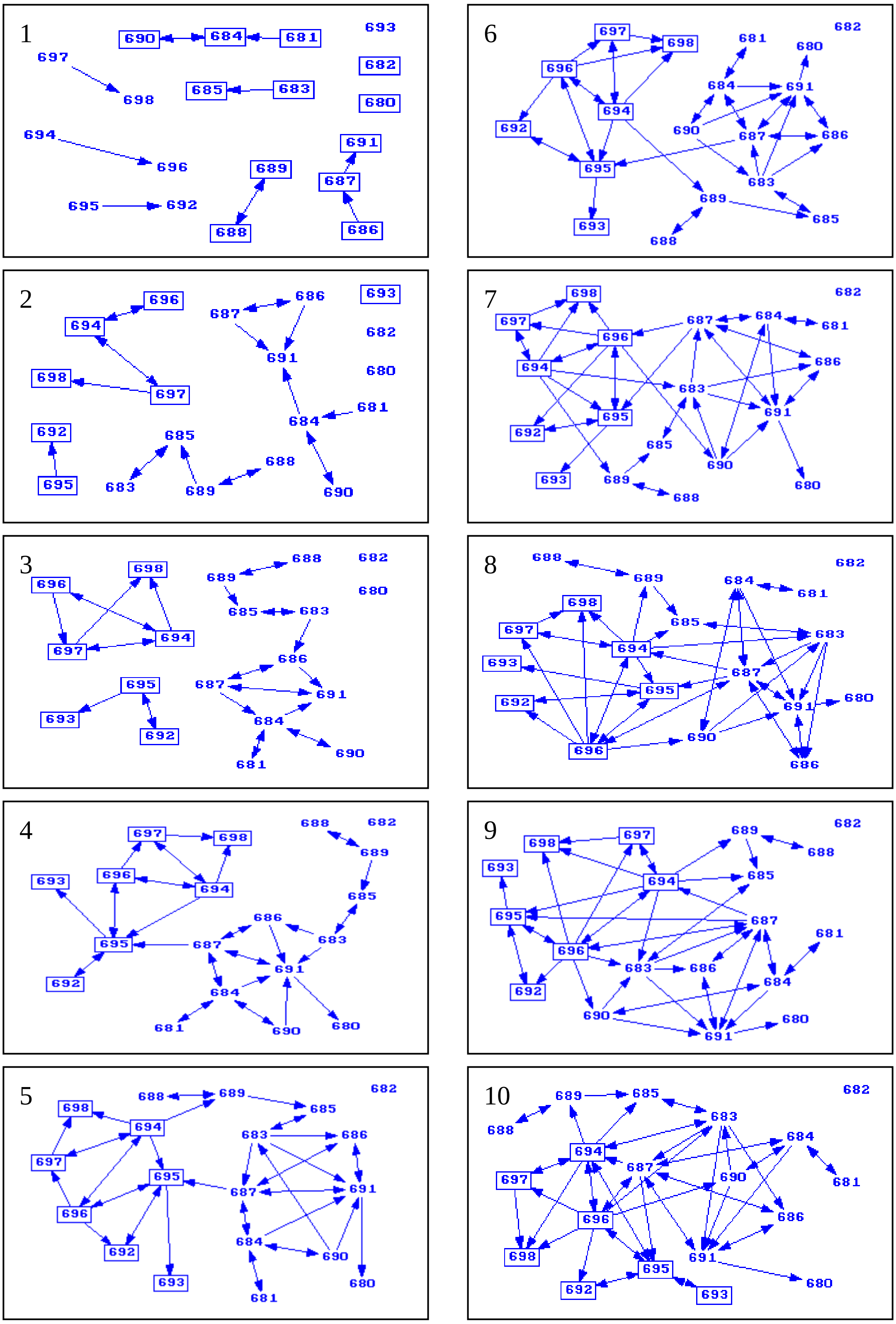
Nomination number	Criteria I No. of class (%)	Criteria II No. of class (%)	Criteria III No. of class (%)
3	0 (0)	1 (2.27)	3 (6.82)
4	1 (2.27)	7 (15.91)	7 (15.91)
5	9 (20.45)	12 (27.27)	17 (38.64)
6	8 (18.18)	10 (22.73)	10 (22.73)
7	6 (13.64)	8 (18.18)	5 (11.36)
8	5 (11.36)	2 (4.55)	2 (4.55)
9	5 (11.36)	3 (6.82)	0 (0)
10	4 (9.09)	1 (2.27)	0 (0)
＞10	6 (13.64)	0 (0)	0 (0)

Note:

Criteria I: Network position was classified as group member, dyad, tree node, isolate type 1, isolate type 2, liaison type 1, and liaison type 2.

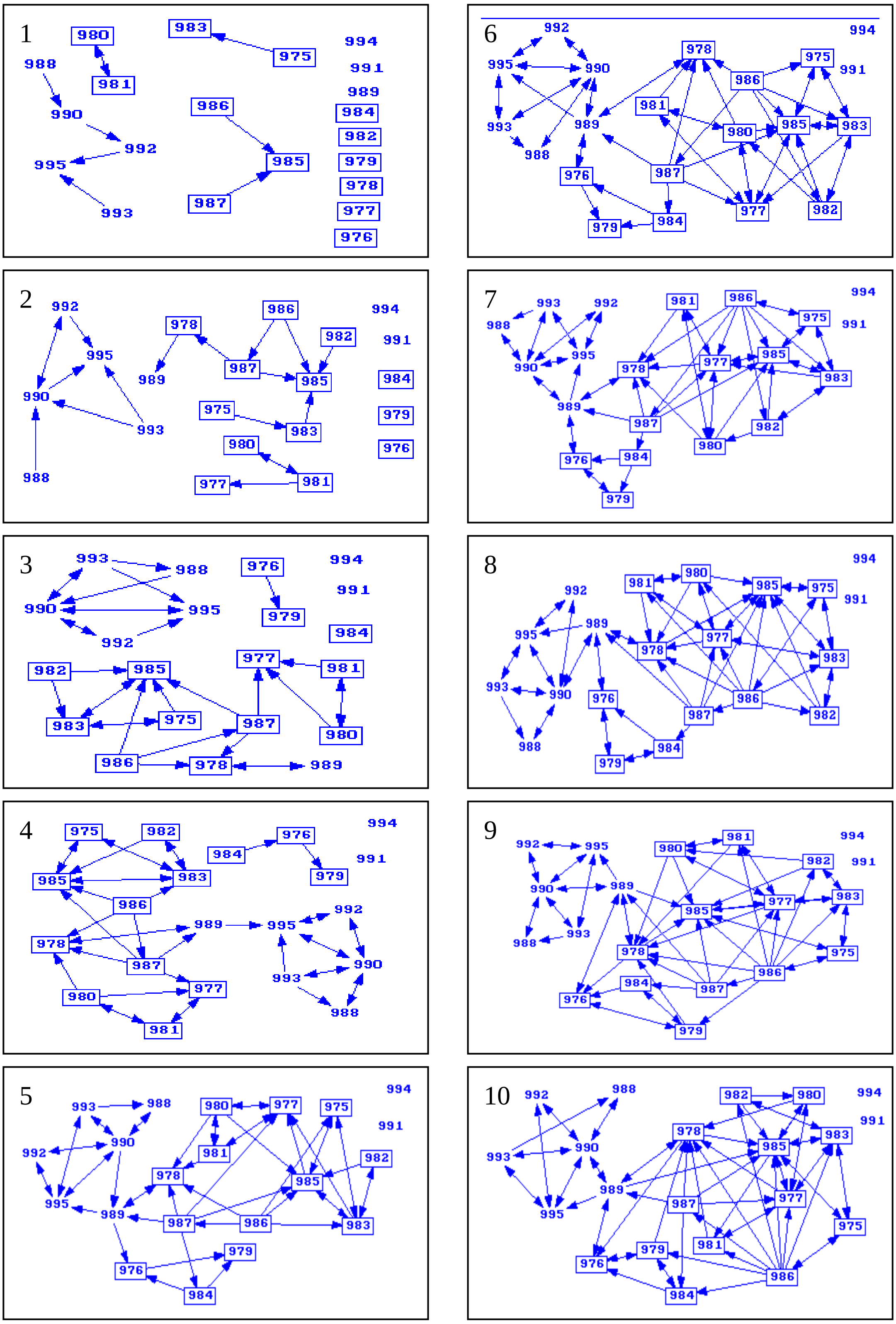
Criteria II: Network position was classified as participants and isolates.

Criteria III: Network position was classified as group members and others.



Note: The number on each network map is nomination number.

Figure 5. Network map of class A (12 boys and 7 girls)



Note: The number on each network map is nomination number.

Figure 6. Network map of class B (8 boys and 13 girls)

References

Aloise-Young, P.A., Graham, J.W., and Hansen, W.B., 1994. Peer influence on smoking initiation during early adolescence: a comparison of group members and group outsiders. J Appl Psychol 79, 281-287.

Bauman, K.E., and Fisher, L.A., 1986. On the measurement of friend behavior in research on friend influence and selection: findings from longitudinal studies of adloescent smoking and drinking. Journal of Yourh and Adolescence 15(4), 345-353.

Borgatti, S.P., Everett, M.G., and Freeman, L.C., 1999. UCINET 5.0 Version 1.00. Natick: Analytic Technologies.

Burt, R.S. 1992. Structural Holes: The social structure of competition. Cambridge: Harvard University Press.

Cairns, R.B., Leung, M-C., Buchanan, L., and Cairns, B.D., 1995. Friendships and social networks in childhood and adolescence: fluidity, reliability, and interrelations. Child Development 66, 1330-1345.

Degenne, A., and Forse, M., 1999. Introducing Social Networks. London: Sage Publication.

Ennett, S.T., Bailey, S.L., and Federman, E.B., 1999. Social network characteristics associated with risky behaviors among runaway and homeless youth. J Health and Social Behavior 40(March), 63-78.

Ennett, S.T., and Bauman, K.E., 1993. Peer group structure and adolescent cigarette smoking: A social network analysis. J Health Soc Behavior 34, 226-236.

Ennett, S.T., and Bauman, K.E., 1994. The Contribution of influence and selection to adolescent peer group homogeneity: The case of adolescent cigarette smoking. J of Personality and Social Psychology 67(4), 653-663.

Ennett, S.T., and Bauman, K.E., 1996. Adolescent social networks: school, demographic, and longitudinal considerations. Journal of Adolescent Research 11(2), 194-215.

Fisher, L.A., and Bauman, K.E., 1988. Influence and selection in the friend-adolescent relationship: findings from studies of adolescent smoking and drinking. Journal of Applied Social Psychology 18(4), 289-314.

Furman, W., 1989. The development of children's social networks. In D. Belle (Ed.), Children's social networks and social supports (pp. 151-172). New York: Wiley.

Haynie, D.L., 2001. Delinquent peers revisited: does network structure matter? AJS 106(4), 1013-57.

Krackhardt, D., Blythe, J., and McGrath, C., 1994. KrackPlot3.0: An improved network drawing program. Connections 17(2), 53-55.

Michell, L., and Amos, A., 1997. Girls, pecking order and smoking. Soc Sci Med 44(12), 1861-1869.

Pearson, M., and Michell, L., 2000. Smoke rings: Social network analysis of friendship groups, smoking and drug-taking. Drugs: Education, Prevention and Policy 7(1), 21-37.

Richards, W.D., 1995. NEGOPY 4.30. Canada: School of Communication, Simon Fraser University.

Scott, J., 2000 Social Network Analysis. 2nd ed. London: Sage Publications.
Urberg, K.A., Degirmencioglu, S.M., and Pilgrim, C., 1997. Close friend and group influence on adolescent cigarette smoking and alcohol use. Developmental Psychology 33(5), 834-844.
Urberg, K.A., Degirmencioglu, S.M., Tolson, J.M., and Halliday-Scher, K., 1995. The structure of adolescent peer networks. Developmental Psychology 31(4), 540-547.
Wasserman, S., and Faust, K., 1994. Social Network Analysis: Methods and Applications. Cambridge University Press.
Yang, H.H., Chen, D.R., Yen, L.L., and Ke, S.R., 2002. The relationship between classmates' interpersonal network and smoking behavior: a preliminary investigation among the 10th graders. Taiwan J Public Health 21(3):164-172.
Yang, H.H., Chen, D.R., Yen, L.L., and Ke, S.R., 2003. Parent and peer group influence on the smoking behavior of high school students. J Med Education 7(2), 128-139.