

# A design to promote group learning in e-learning: Experiences from the field

Pei-Chen Sun <sup>a,\*</sup>, Hsing Kenny Cheng <sup>b</sup>, Tung-Chin Lin <sup>c</sup>, Feng-Sheng Wang <sup>c</sup>

<sup>a</sup> *Institute of Information and Computer Education, National Kaohsiung Normal University, 116 Ho-Ping First Road,  
Kaohsiung 802, Taiwan, ROC*

<sup>b</sup> *Department of Decision and Information Sciences, Warrington College of Business Administration,  
University of Florida, Gainesville, FL 32611-7169, USA*

<sup>c</sup> *Department of Management Information System, National Sun Yat-Sen University, 70 Lien-hai Road Kaohsiung 804, Taiwan, ROC*

Received 4 April 2006; received in revised form 19 July 2006; accepted 19 July 2006

---

## Abstract

The Internet enables learners to be brought together where they can cooperate in learning in groups without space and time limitations. It is, however, quite a challenge to form ideal groups in a short time and ensure satisfactory interaction for learners in cyberspace. In this study, we propose a useful grouping method to help teachers improve group-learning in e-learning by first establishing effective groups with rules based on data mining, and then facilitating student interaction using a system that monitors members' communication status. Field observations and quantitative evidence show the validity and practicability of the proposed method.

© 2006 Elsevier Ltd. All rights reserved.

**Keywords:** Group learning; e-Learning; Data mining application

---

## 1. Introduction

Recent computer-based instruction (CBI) pedagogies embrace the Internet to enhance group learning performance. While there has been much research on group learning, little attention is paid to a rather practical and yet critical aspect of group learning in the setting of e-learning involving Internet – how to form in a minimal amount of time highly interactive groups that achieve effective learner to learner and learner to instructor communications. Such a practical and critical issue is significant for group ecology (Brush, 1998; Kirschner, 2001). As group learning on the Internet becomes more and more popular for modern curricula of higher education, the need has never been greater for instructors to find an innovative way to form highly interactive e-learning groups (Scott, Chanidprapa, Seung, Jared, & Jason, 2002).

---

\* Corresponding author. Tel.: +886 7 717 2930; fax: +886 7 7117529.  
E-mail address: [sun@nkn.edu.tw](mailto:sun@nkn.edu.tw) (P.-C. Sun).

Several researchers suggest that group interaction is a vital element in team success (Lebie, Rhoades, & McGrath, 1996; Wanger, 1995). Particularly in distance education and e-learning, researchers note that information exchange and interaction with others enhances student performance and bolsters high satisfaction levels (Gear, Vince, Read, & Minkes, 2003; Vrasidas & McIsaac, 1999). The opportunities for productive interaction during online learning, unfortunately, are scarce even though the importance of interaction has been well established in the literature. In general, online learning affords fewer opportunities for non-verbal feedback, which is prevalent in face-to-face communication (Hara, Bonk, & Angeli, 2000; Jonassen & Kwon, 2001; Pena-Shaff, Martin, & Gay, 2001). Further, when people work collaboratively but not face-to-face, many interaction resources are disrupted. Some research suggests that it is important to develop interaction strategies in these computer-mediated environments (Pena-Shaff et al., 2001; Polichar & Bagwell, 2000). However, most e-learning courses at present lack group coherency and identity that positively impacts information exchange and knowledge sharing especially in a problem-solving scenario.

Instructors of online courses usually find that their class size is larger than that of traditional classes and students rarely interact effectively. Forming groups with high levels of interaction is difficult for most students because of the limitations of social intercommunication on the Internet. Researchers devoted to establishing an e-learning platform for universities also face the same challenge. Little extant literature focuses on how to form online groups that achieve efficacies of time and interaction. Both researchers and instructors are in dire needs of overcoming this predicament arising in e-learning.

In this study, we propose a method that helps instructors to form high-interaction e-learning groups by applying rules based on data mining techniques. Further, to facilitate student interaction, we develop a system that monitors group members' communication and provides valuable group interaction information to the instructors. Using the grouping method and the monitoring system we develop, instructors can form high-interaction groups efficiently, and they can facilitate the group interaction processes on the Internet. The effectiveness of our grouping method and online monitoring system is validated through field experiments. The overall design of our system is shown in Fig. 1.

The rest of this paper is structured as follows. Section 2 describes how to apply data mining techniques to discover rules for high-interaction groups. In Section 3, we propose the use of clustering techniques to help create effective groups during e-learning. Section 4 presents the research methodology and the facilitation mechanisms applied in this study. The main results and discussions drawn from the experiments are presented in Section 5. Section 6 provides concluding remarks, practical implications of our research, and future work.

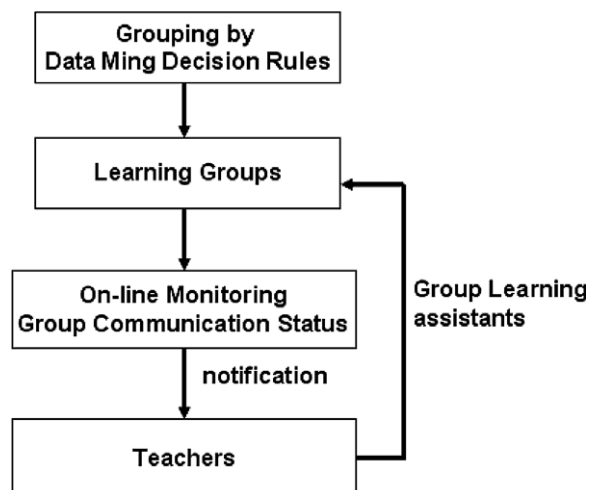


Fig. 1. Design for effective groupings in e-learning.

## 2. Data mining from high-interaction groups

Communication and information exchange among group members is found to be the most important action with partners to complete group goals (Johnson & Johnson, 2000; Webb & Palincsar, 1996). Numerous studies find that frequent and accurate interaction among group members results from appropriate matches of group members' attributes, including internal and external ones (Lee, 1993; Webb, 1982).

In traditional instruction, students find their favorite or suitable members to form a group for teamwork. This form of grouping is considered better for interaction but difficult to implement in the e-learning setting since students taking the online courses generally do not know each other and lack the face-to-face contact to "feel" out potential group members.

If we can identify the rules of choosing members for high interaction within groups, applying these rules will ensure group interaction in e-learning. In this section, we describe data mining and clustering techniques to discover rules and relations that unveil the attributes of high-performance groups from 415 students taking a web-based course in Fall 2002. In particular, the rules of most important attributes affecting the satisfaction and interaction of group members were discovered by applying the ID3-C4.5 classifier (Quinlan, 1993).

### 2.1. Attribute selection

Several researchers suggest that group design elements are critical for social interaction and group success, while others maintain that individual differences are central (Edwards, 1996; Marwell & Ames, 1979; Wanger, 1995). Interaction among members and their relations inside groups can be regarded as revealing preferences that relate to collocations among certain internal and external personal attributes, such as gender, age, place of residence, preferences, interests, and values (Marwell & Ames, 1979; Wanger, 1995). A critical question is what attributes of group members can be used to guide the forming of high-interaction groups? Studies suggest that group members of web-based learning should be allowed to interact concurrently with few barriers to cooperation (Riel & Harasim, 1994). Ellis and Fisher (1994) find that personal characteristics will affect the communication modes within the group learning process. Evidently, human factors such as the attributes of people are of considerable importance for members when they are to form an organizational group.

Many empirical studies indicate that successful groups always have a good match of members' attributes (Baer, 2003; Mitchell, Rosemary, Bramwell, Lilly, & Solnosky, 2004). We consider both internal and external attributes when integrating attributes and constructing attribute sets to complete group profiles. Based on the prior literature and the particular e-learning setting at hand, we identify four different attributes for further research of high-interaction and high-satisfaction groups – learning periods (Time), Region, Age, and Value types. These four attributes of 415 students taking the web course held in Fall 2002 were collected from the database records and pre-questionnaires over the 14-week period of the course. The 415 students were requested to form learning groups of five individuals in each group. Some of them formed groups on their own, while others did not.

#### 2.1.1. Learning periods (time)

We find that students taking online courses via the web have their favorite learning periods, interact with one another at specific times, and make arrangements to access information simultaneously, implying preferences in working times and group norms. Based on the disciplines of interaction described in the Information Richness Theory (Daft & Lengel, 1986) and the Social Influence Theory (Fulk & Boyd, 1991), there will be similar patterns of media use among group members. Media choice is at least in part a group norm, a regularity of group behavior that emerges for reasons outside the optimization of individual or group utility (Fulk, Schmitz, & Steinfield, 1990). Interaction will be established with the idea that people can enjoy access to informative communication resources at the same time (Daft & Lengel, 1986). According to the records of past e-learning courses at the authors' universities, students have several different favorite periods during the week to learn on the web. We capture students' login times for the web-based course under consideration and find three periods of time with higher login frequency (working days at 8:00–11:00 PM and weekends at 9:00–12:00 AM and 2:00–5:00 PM) to be the times of contact preferred by most students.

### 2.1.2. Regions

In an asynchronous web-based learning environment, people come from different regions with different customs, cultures, and cognitive modes, and will have specific preferences in choosing partners to make up social networks in groups, which may influence their interactions with and acceptance of those from other regions (Blaszczak, 1990). Therefore, regional effects exist when people select group members for online courses. It seems that regions influence human cooperation in a wider environment where regional culture and social content affect attitudes of human behavior (Middleton, 1976; Tuch, 1987). From our past experience of grouping students taking Web courses, students regarded regional difference as an important consideration for cooperation and collaboration in teams of e-learning. For this reason, our grouping system examined the regions of those of the 415 students who were thought to be most active in interactions. Four geographical regions of residence making up the region attribute include North, Middle, South and East.

### 2.1.3. Ages

There are differences between generations, for example, in attitudes, lifestyles, and economic interests. Longer and lasting cooperation between members of different age groups (on condition of equality) and joint success should reduce intergenerational conflicts and, therefore, improve attitudes among members of different age groups (Pinquart, Wenzel, & Sorensen, 2000). Studies find that different age intervals and age gaps will bring diverse reliability and identification in organizations (Gilbert & Tang, 1998). From many empirical studies (e.g., Holladay & Kerns, 1997) and our observations on group learning, “age” is a significant factor that influences partnerships among group members. It has been shown in traditional learning that age has significant effects on people selecting group members. Likewise, the age attribute should affect student groups’ satisfaction in web-based learning. After capturing the age data from 415 students of the aforementioned web course, we divide the attribute of age into six groups based on intervals of five years – age 23–27 (group-25), age 28–32 (group-30), age 33–37 (group-35), age 38–42 (group-40), age 43–47 (group-45), and age 48–52 (group-50).

### 2.1.4. Value types

The last, but certainly not the least, attribute used for grouping students taking online courses into high-interaction groups is value type. The starting point for choosing the value type attribute is on the basis of its relativity, not its absoluteness, since value type has a neutral nature and won’t result in the so-called “bad group” or “good group” matching (Allport, Vernon, & Lindzey, 1960). Allport et al. (1960)’s *A Study of Values* aims to measure the relative prominence of value whose classifications is based directly on Eduard Spranger’s *Types of Men* (Spranger, 1928), a brilliant work which defends the view that personalities of men are best known through a study of their values or evaluative attitudes. People can be classified into one of the six value types that people manifest from the value scale and manual in *A Study of Values*. The value scale in *A Study of Values* has been used in testing personalities for over 70 years, and its reliability and validation are both strongly proven in anthropology, psychology and other social sciences. Therefore, we developed a questionnaire based on Allport’s *A Study of values* to survey personal data from 415 students in the web-based course of Fall 2002. We have successfully classified each subject into one of the six value types (theoretical, economic, aesthetic, social, political, and religious).

## 2.2. Learning group’s attributes

There were five students in each group for web-based learning and each student was observed in terms of the four attributes described in Section 2.1. The specific pattern of each attribute collocation can be initially considered in terms of how different it is from the other attributes in the group. We expect the high-performance groups to have similar attribute patterns. The designation, “similar attributes,” can serve as a sign of “higher approximation” among the instances for each attribute. Since a simple approximation function can measure multiple vectors, we design the approximation of the five instances for each attribute of a group.

We design the approximation function to calculate group-attribute simulation before inputting the attributes under consideration into the ID3 classifier. Take the attribute of “Ages” as an example. If all five members of a group have the same age attributes, we can record this combination as “1,1,1,1,1” (AAAAA or BBBBBB, and so

Table 1  
Classification of three interaction levels among web-learning groups

	Number of groups (total: 83)	Average satisfaction response	Synchronous (chat room)		Asynchronous (discussion forum)	
			Group (h/week)	Single (h/week)	Group (post/week)	Single (post/week)
High level	23	5.33	3	2.5	8	1.6
Normal level	38	4	1.5	1	4	0.8
Low level	22	2.66	1	0.5	2	0.4

on). We choose this frequency type as a seed to compare its approximation with other types, and its group-attribute simulation is set as 4 to serve as a “flag” for measuring other groups. However, if one member of the group shows a different attribute value, we record its combination as “1,1,1,1,2” (AAAAB or BBBBC or CCCCD, and so on), and its group-attribute simulation as 3. If this group has three members having the same attribute value with the rest having the same attribute value but different from that of the other three, we code its combination as “1,1,1,2,2” (AAABB or BBBC or CCCDD and so on) and its group-attribute simulation as 2.5. The other three attributes – Times, Region, and Value type – are computed in the same way.

### 2.3. Research data

There were 415 students enrolled in a 14-week course at a university in Taiwan, R.O.C. in Fall 2002. All students were asked to form learning groups of five members. They were required to login to a Web-based Instruction (WBI) system to study and finish a team task by the 10th week. Their communication data were captured by the WBI system, including the frequency of interaction in the discussion forum and chat room. We found three distinct interaction levels from those 415 students as described in Table 1. A questionnaire investigating group satisfaction was answered by each group and returned to the instructor. By the 8th week, all data regarding attributes and satisfaction level were analyzed. Twenty-three groups had average satisfaction level more than or equal to 5.33, a Q3 value considered to be a high satisfaction point. Twenty-two groups showed their average satisfaction level to be less than 2.66. These groups with lower satisfaction were categorized as “Low S”. Thus, the Q1 value of 2.66 was considered to exhibit low satisfaction. The two opposing extremes of satisfaction were designated High S and Low S.

All of the four attributes data in the 83 groups had to be classified to discover rules about satisfaction. It was necessary to simplify the ranking of group-attribute simulation to make calculating and rule-mining easier. After a cursory observation of all group-attribute simulation for the 83 groups, a heuristic bracket was devised by statistical distribution. When each attribute set was examined, the 60 percentile point among the approximation values (2.5) (Time: 60.6%, Region: 59.35%, Age: 57.8%, Value: 58.05%) was shown to have good discriminating ability for the sample data. Therefore, the group-attribute simulation values can be redefined dualistically for simplification. For this reason, we decided that if group-attribute simulation value was higher than 2.5, it would be recorded as High \*Si; otherwise, the lower value would be recorded as Low \*Si.

### 2.4. Mining results

The Expert Rule is a data mining tool noted for its great capability and rapid outcome as the ID3-C4.5 kernel. Attribute data and the two interaction levels (high and low) of 45 groups were entered for classification in order to extract the relations between attributes and their heuristic rules.<sup>1</sup> For easier understanding of the detailed records of classifying all observed groups, the results of the Expert Rule were transformed into a visual illustration of rules classification in Fig. 2. All groups were well classified in the final sets by the Expert Rule Classifier. After observing the decision rule paths, we identified four rules of forming high-interaction groups. Fig. 3 shows the attribute path of each rule resulting in high-interaction groups.

<sup>1</sup> Since we wanted to find out the rules distinguishing high satisfaction groups from low satisfaction ones, 38 groups with average satisfaction were excluded from the classification. Therefore 45 rather than the all 83 groups were entered for classification.

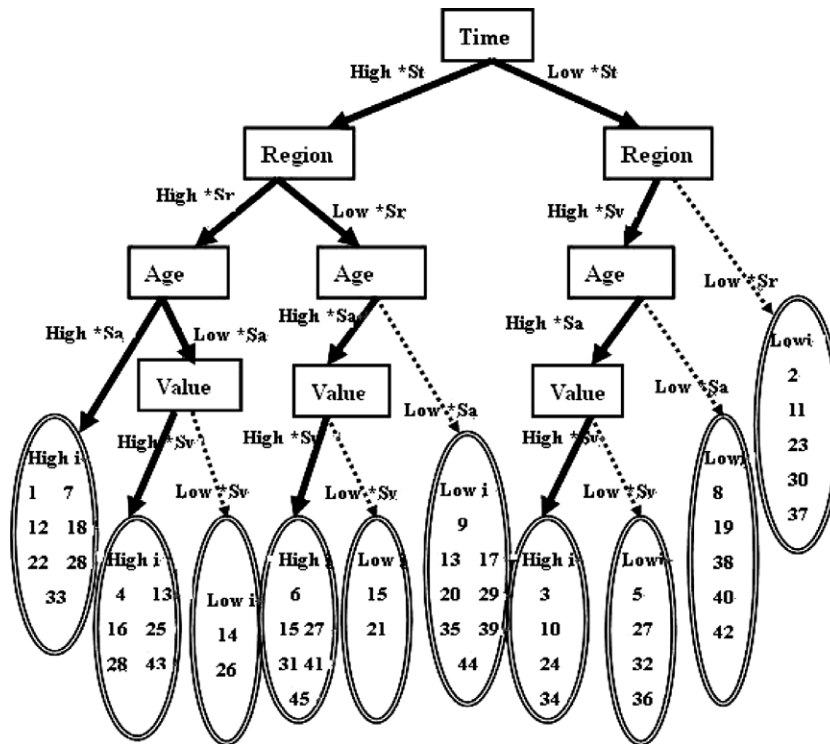


Fig. 2. Classification of rules of groups.

**R1** → Low \*St(Time) ∧ Low \*Sr(Region) ∧ Low \*Sa(Age)

**R2** → Low \*St(Time) ∧ Low \*Sr(Region) ∧ High \*Sa(Age) ∧ Low \*Sv(Value)

**R3** → Low \*St(Time) ∧ High \*Sr(Region) ∧ Low \*Sa(Age) ∧ Low \*Sv(Value)

**R4** → High \*St(Time) ∧ Low \*Sr(Region) ∧ Low \*Sa(Age) ∧ Low \*Sv(Value)

Fig. 3. Four rules of high-interaction groups.

Three findings are discerned from the data mining results of the Expert Rule. First, 20 of the 23 groups with high interaction were formed based on their own preferences as to choice of group members, which unsurprisingly resulting in high satisfaction. Empirical studies show that if people can choose their preferred group members, they will interact more readily and freely. Eighty-seven percent of the groups in our survey formed by choosing their own members enjoyed a high level of interaction, thereby confirming the known empirical results. Second, one can easily see from Fig. 1 that the initial root-node is Time, the learning periods of the groups. This means that the first critical attribute of each rule is Time, which serves as a starting point of good classification. In other words, the learning period is a strong attribute, which can heavily influence the interaction in the first stage of a web-based group. When people in a group have the same learning periods on the Web, they are bound to have more opportunities to interact and to experience more cooperation. Third, analyses of R1, R2, R3, and R4 for their characteristics generalized via the decision tree seem to suggest that each rule has a high proportion of High \*Si attributes (R1: 1, R2: 3/4, R3: 3/4, R4: 3/4). After checking all states of attributes for each rule, we can develop a heuristic hypothesis as “if group members in e-learning have many similar attributes, they will interact more productively as a group”.



### 3. Forming groups by clustering techniques

In the previous section, we used the group-attribute simulation value (\*Si) as a group's representative attribute. The statistical indicators of group-attribute simulation can be switched to explain the dispersed or centralized appearance of the samples. High \*Si means that individuals are similar to each other in terms of less different values of their characteristics. Fortunately, it is an opportune coincidence that clustering techniques can accomplish this type of tasks. The main methodology of data clustering is to classify data by the principle of the highest number of similar attributes, and then to distinguish between more similar and less different attribute groups (Groth, 1997). The clustering algorithm can be easily adapted to design a systematic grouping method in our e-learning setting.

Let  $n$  objects (i.e., students) be described by their respective attribute vectors  $\{X_1, X_2, X_3, \dots, X_n\}$ . Each vector contains the values of a student's attributes. For example, if the  $j$ th object is a 35 years old student taking the 9 a.m.–12 noon class from the north region with the economic value type, then the corresponding attribute vector of  $X_j = \langle \text{Time, Region, Age, Value Type} \rangle$  has the instantiation of  $\langle 9 \text{ a.m.} - 12 \text{ noon, north, 35, economic} \rangle$ . Research shows that the Euclidean distance between objects  $X_1$  and  $X_2$  can be used to measure their difference. Let  $c_i$  be the mean of the vectors in cluster  $i$ . That implies that object  $X$  is in cluster  $i$  if the distance between  $X$  and  $c_i$  is the minimum (Dash, Liu, & Xu, 2001). Our algorithm in this study refines the nearest-group-method, which is a discriminate analyzing and classifying approach in statistics. The measure scale uses the approximation function for calculation to assemble close data into a cluster. A required condition of grouping students into clusters in this study is to maintain a more appropriate group size. Hence, an adjustment mechanism was added to our algorithm to keep the differences minimal and the group size ideal. Our algorithm is described as follows:

1. Start with  $n$  clusters with a single object in each cluster.
2. Find the most similar clusters and merge them into one cluster.
3. Repeat Step 2 until the total number of clusters is 1 and the hierarchical cluster tree is obtained.
4. Based on the hierarchical cluster tree derived in Step 3, start grouping with four to six objects in each group in a bottom-up fashion.

#### 3.1. Grouping implementation

We implemented our grouping algorithm and conducted a field experiment to validate our research in a real instructional environment. We chose a Web-based course to conduct our experiment based on the instructor's willingness and its large amount of students. There were 202 students taking this web-based course over 14-weeks of learning sessions in Spring 2003. All students were assigned into one of two different types of groups – either formed by our grouping algorithm, hereafter denoted as “system” groups, or formed at random, hereafter termed “random” groups, which served as the control sample. The same group tasks and weekly assignments were assigned to all groups. Each group had to brainstorm and hold discussions through chat-room of the WBI system. All pertinent data were captured during this process.

#### 3.2. Research sample and course schedule

A pre-experimental design similar to that described in Section 2 was carried out before the actual field experiment. In this field experiment, the 202 students were separated into either system groups or random groups. We chose 100 students to be our experiment subjects who were to be grouped into 20 system groups by our grouping algorithm and monitored during the whole e-learning process. The other 102 students were randomly assigned into 20 random groups and not monitored while they were learning on the web. Students had 3 weeks to become familiar with each other on the Internet, and then form groups either by our grouping algorithm or randomly as the 4th week began. Students had 3 weeks to become familiar with the WBI system and to freely engage in social and cyber interactions. Having been assigned weekly assignments, all groups were asked to proceed with their group discussions when the 4th week began. All students were to finish their

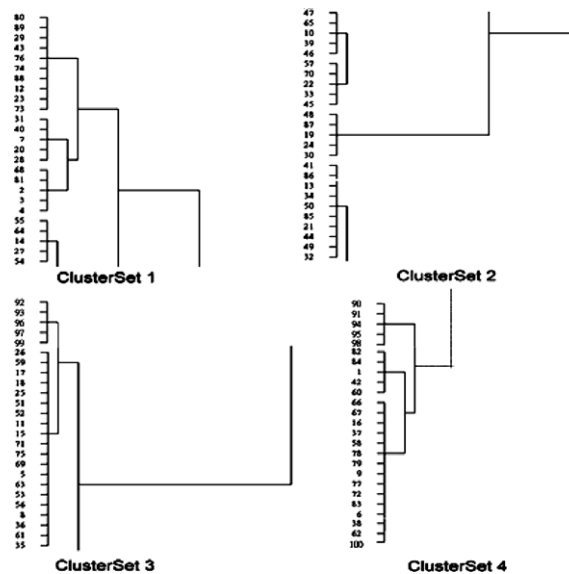


Fig. 4. Twenty groups created by clustering.

Table 2  
Members of group by clustering algorithm

Cluster	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Members' serial numbers	1	2	5	6	7	8	9	10	11	12	13	14	16	17	19	21	22	29	90	92
	42	3	53	38	20	35	72	39	15	23	34	27	37	18	24	32	33	43	91	93
	60	4	63	62	28	36	77	46	51	73	41	54	58	25	30	44	45	76	94	96
	82	68	69	83	31	56	78	47	52	74	50	55	66	26	48	49	57	80	95	97
	84	81	75	100	40	61	79	65	71	88	86	64	67	59	87	85	70	89	98	99

group projects by the 9th week, and to hand in their assignments in the 10th week. All weekly assignments were addressed by the instructor and each weekly assignment was highly correlated to the group project to be handed in during 10th week. All groups had 6 weeks to interact and collaborate in order to finish the same group project. To verify students' learning effects, a satisfaction questionnaire was mailed to each group with the stipulation that it was to be sent back along with their group projects.

### 3.3. Cluster-grouping

Before applying the clustering algorithm to group those 100 students selected for our experiment, the four attributes described in Section 2.4 – Time, Area, Age, and Value – had to be numerically encoded and entered into the algorithm. The encoding number of each attribute and their descriptions were designed exactly in the same way as described in Section 2.1. All of the encoding data were computed via group-attribute simulation as the distance by the clustering algorithm. We found that those 100 students were clustered by our algorithms very well as shown in Fig. 4. Table 2 shows the detailed makeup of each system group generated by our clustering algorithm.

## 4. Design of the “interaction monitor”

As the 4th week began, there were 20 system groups and 20 random groups ready to take on group assignments. All of them were asked to interact via the group-discussion area in the WBI system. The interactions among members of each of the 20 system groups were closely monitored by an OLAP-based kernel called the



On-line “interaction monitor” module embedded in the WBI system. When students logged onto the WBI system and engaged in group interactions, this module captured data related to those interactions. To effectively help instructors understand the ever-changing conditions of group interactions, this module analyzed several critical indicators and notified instructors by e-mails twice a week. After an instructor received the notification e-mail and linked to the module through password authentication, he or she could get a quick overview of group interaction. The data provided by the “interaction monitor” were considered very helpful by instructors to know how groups were interacting and to identify which group or student was facing interaction barriers.

There are four major parts of the notification information from the on-line “interaction monitor”. The first part displays the summary information of the current interaction sessions of all groups (see Fig. 5), which distinguishes between the better-interaction groups and the worse-interaction groups. We set the interaction sequences at 3.5 days a round to analyze the interaction levels of all groups. The average group and its scores for several important indicators are computed and shown in the summary information window. According to the characteristics of group interactions on this WBI system, we often choose four indicators, average usage time on group-discussion, the amounts of messages sent/received, the content bytes of the messages, and the times of successful interaction from member to member. Each indicator is closely observed and analyzed by the “interaction monitor”, and the interaction performance of each group is ranked as either better-interaction or worse-interaction. Instructors can also check the historical data of all groups to judge the ranking of different sessions.

The second part places interaction records of specific group for the instructor’s perusal, see Fig. 6. For example, suppose an instructor would like to know why group No. 13’s interaction was below par in the current session. He or she can then review the analysis of the interaction-indicators and the detailed interaction activities. We include a visual diagram of interaction linkages to clearly pinpoint the level of interaction, see the right pane of the Fig. 6. In this pane, all detailed successful interactions and their linkages are recorded and displayed through the topography of group interaction with nodes and lines. Based on the number of successful interactions, diverse lines and colors show each interaction level from node to node. Instructors can easily recognize who urgently needs assistance by this diagram and the status of member ranking by counting all interaction linkages.

The third part of the interaction monitor is devoted to setting observation parameters (see Fig. 7), including basic operations, indicator selection, notification modes, path of historical-tracking, and so on. Instructors can select options for how to represent information, change the plan for monitoring specific indicators and organize personal settings to match their instructional strategies. We design a default status for all settings and suggest to instructors ways to adapt the settings when using the system for the first time. Depending on his or her pedagogical discipline, the instructor can also configure advanced settings for detailed functions.

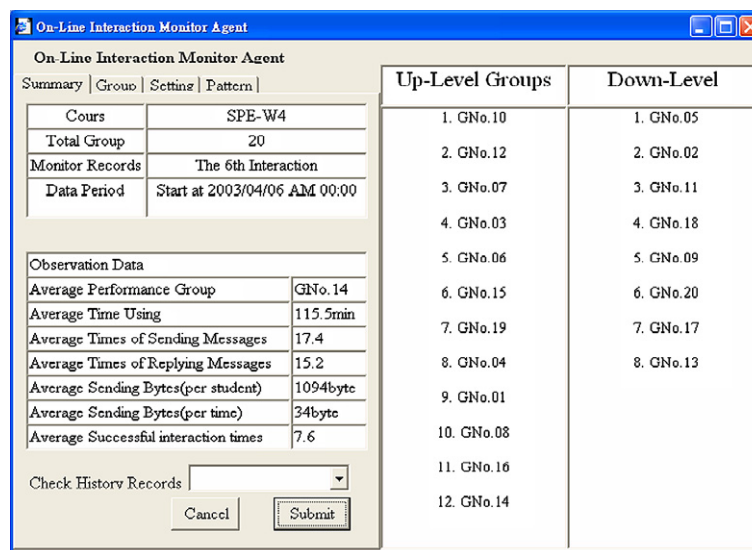


Fig. 5. Summary information page of interaction monitor.

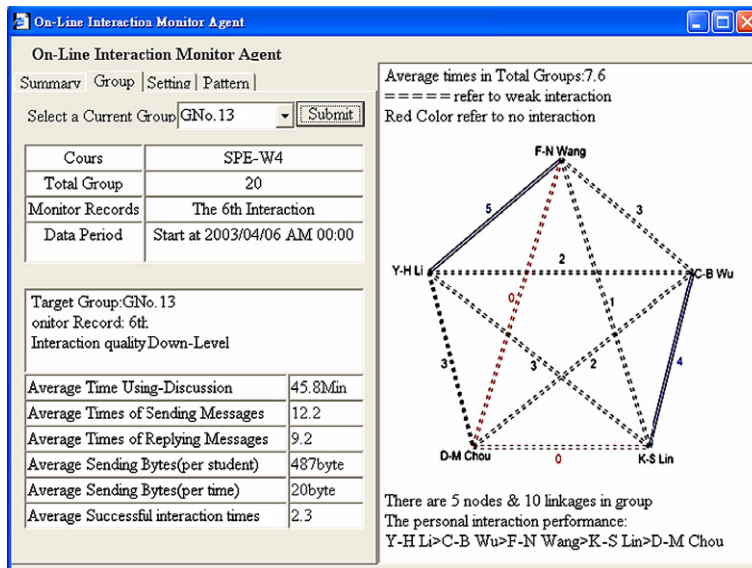


Fig. 6. Interaction records of specific group.

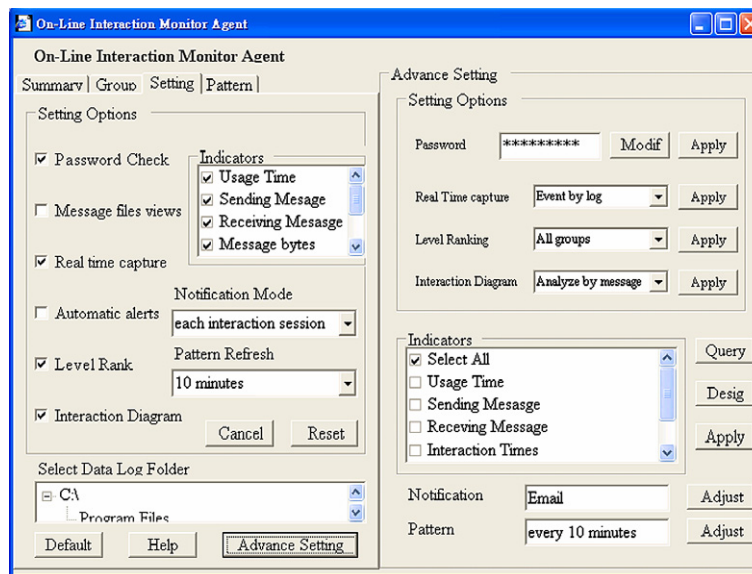


Fig. 7. Setting function of optional observation.

The final part is the pattern analysis of group interaction (see Fig. 8) where instructors can view the trends of certain indicators by group or by time. According to the conditions set by the instructor, the interaction monitor module will analyze data and load clear sketches. Instructors can set three reference conditions (interaction-session period, specific group, shape style) and reconfigure resultant diagrams as a bar chart, polygon, or trend curve, etc. To further understand the variations among groups or different periods, instructors can choose to execute historical comparisons. The interaction monitor includes a tool program that links to more powerful analysis software such as MATLAB.

To evaluate the efficacy of our grouping approach and interaction monitor, two hypotheses are proposed as follows:

*Hypothesis 1.* Groups that are formed using the proposed grouping algorithm and with interaction monitor are more effective groups than random groups without interaction monitor.

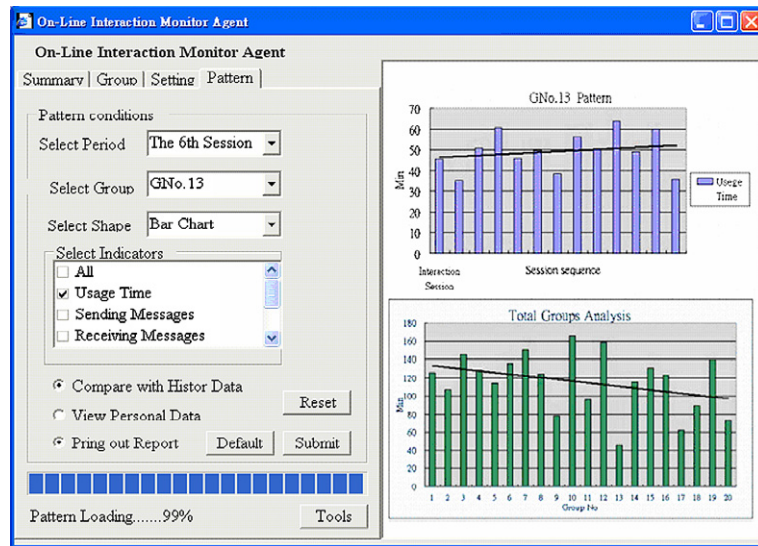


Fig. 8. Pattern analysis function of interaction monitor.

*Hypothesis 2.* Groups that are formed using the proposed grouping algorithm and with interaction monitor will spend more time on discussions in WBI than random groups without interaction monitor.

## 5. Research results

To validate the beneficial effects of our method of grouping and interaction facilitation, we undertook both a qualitative method of interviewing the instructors and students and a quantitative method of statistical analyses of data from the questionnaires and the interaction monitor since multiple data collection methods provide a more complete picture of the research issues (Pinsonneault & Kraemer, 1993). After the final due date of the group tasks in the 10th week, we interviewed the instructor of the Web-based course and asked what students had learned in this course. The instructor gave very positive response to our grouping method and interaction facilitation function provided by our system as follows:

“Compared with other WBI courses in the past, I benefited greatly from the system this time. The assistance of grouping and notifications were most convenient. Those groups formed by the system were more devoted to group activities. I felt they could quickly develop their team’s centripetal force. At the same time, I saved a lot of time by not having to pay close attention to every group. I just concentrated on certain groups that interacted poorly according to notifications from the system. I could quickly and clearly identify which groups had interaction problems and what problems were occurring. I could direct my effort to specific groups or individuals and modify my strategies toward them without wasting as much effort as before.”

Obviously, the instructor offered a positive recommendation of our grouping method and interaction facilitation tools. He stated that our system helped him improve the quality of instruction. Indeed, we received strong encouraging responses from the instructor; but how did the students feel? To gather the opinions and preferences of students in this course, we selected for interview six students who served as group leaders, with half of them from the random groups and the other three from system groups. Each group leader had five minutes to express his or her perception of the grouping process and interaction activities. The students’ responses in below are identified by the type of students with SG for system group and RG for random group, followed by more specific identification.

“I am satisfied with my partners because I feel they became familiar with me quickly. We also interacted well and shared much time in discussion.” (SG-GNO.10 F-N Lin) “We seldom suffered from severe learning barriers in this course. There were lots of interactions between members, and we had many chances to meet in the chat-room at the same time.” (SG-GNO.03 G-H Ou) “Our teacher would send messages to

certain members when they had fewer opinions during discussion or when they missed the discussion. Our instructor often would join our discussion or assist someone who felt it strange to share his or her thoughts. In general, we formed a good consensus in completing our group task.” (SG-GNO.08 L-D Wang)

Not surprisingly, students of the RG type had quite different impressions of their learning process as follows:

“There were some troubles in my group. I feel it took too much time to get to know my partners. We even had less similar topics of conversation on the course.” (SG-GNO.13 Y-M Li) “Because of the absence of cohesion in learning, we didn’t interact frequently or even talk much. Sometimes, when I posed a question, nobody would be willing to answer me soon.” (SG-GNO.18 K-M Chio) “Unfortunately, we lacked stimulation in our interaction and people felt it strange to share thoughts. I think we really interacted badly during discussion about the group task.” (SG-GNO.15 R-P Huang)

From the messages students revealed above, we find three important issues that need to be examined – (1) the satisfaction levels of students within these two different types of groups, (2) the effects of their usage and efficiency of interaction, and (3) how different groups interacted at different stages of learning. To address the first issue, we used a questionnaire to collect students’ satisfaction levels during the group-learning process. The remaining two issues were analyzed by extracting data from our WBI system.

Since the questionnaires were to be sent back along with the group task results, the response rate was 100% and there were 40 valid questionnaires. The satisfaction questionnaire was designed by Gladstein (1984) to investigate the level of team satisfaction. Three items were developed to survey team satisfaction with the Likert 7-point scale (e.g. 1 = strongly disagree, 7 = strongly agree). After analyzing all questionnaires, we found the Cronbach- $\alpha$  value for system groups to be 0.9429, and 0.9238 for random groups, far exceeding the recommended reliability threshold of 0.7. This result suggests that our questionnaire was very reliable. The descriptive statistics are shown in Table 3.

According to Table 3, the satisfaction levels of random groups were lower than those of system groups. In order to find whether the satisfaction of random groups is significantly different than that of system groups, one-way ANOVA was adopted to test the level of significance and  $F$  value, see Table 4.

Table 4 shows that indeed system-grouped students’ satisfaction during group learning was quite higher than that of random groups at the 0.05 significant level ( $p^* = 0.001$ ).

Students of system groups were satisfied with their fellow group members, and they were also assisted and prompted by the instructor aided by the notifications of our interaction monitor, which were the two distinguishing features that students of random groups sorely lack. In terms of usage time of the WBI system, we postulate that students of system groups will spend more time on discussions in the WBI during the learning process. They will show more enthusiasm and interests in discussing with their partners, and they tend to do so even more when the instructor intervenes. To examine the differences in usage time of the WBI system between system groups and random groups, we used the one-way ANOVA to test the significance and  $F$  value, see Table 5.

Table 3  
Descriptive statistics of team satisfaction

Grouping	# of groups	Average	SD	SE
System	20	5.3924	1.129	.2341
Random	20	3.2561	.8263	.2287

Table 4  
One-way ANOVA table of satisfaction

	SS	DF	MS	$F^*$	Significant
SSC	29.48	1	29.48	14.88	.001***
SSE	75.64	38	1.98		
Total	61.12	39			

Table 5  
One-way ANOVA table of usage time

	SS	DF	MS	$F^*$	Significant
SSC	98.06	1	98.06	24.6	.000
SSE	150.45	38	3.9		

Table 5 shows that students' engagement in group discussions among system groups was markedly higher than that of random groups at the 0.05 significant level ( $p^* = 0.000$ ). Selective intervention from the instructor during interaction further promoted the discussions among students of system groups as they learned on the WBI system. The instructor's intervention drove students to become more motivated and willing to devote more effort to discussion.

Since our interaction monitor module enabled instructors to modify intervention strategies when certain groups or certain students had problems with interaction, we expect this facilitation mechanism to raise the success rates of interaction defined as consummated communications through impulse and feedback between the message sender and message receiver. To examine the difference in the success rate of interaction between the two different types of groups, the one-way ANOVA was also adopted to test the level of significance and  $F$  value, see Table 6.

Table 6 shows that interaction success rates of group discussion within system groups are quite higher than those of random groups at the 0.05 significant level ( $p^* = 0.000$ ). On-line interaction monitoring of interaction and facilitation from the instructor increases the quality of discussion and improves the rate of successful interaction for students of system groups much more so than those of random groups.

We have shown that groups formed by our clustering algorithm perform much better overall than groups formed at random in terms of group satisfaction, usage time of WBI system, and successful interactions among group members. Another interesting question is how these two different types of groups perform at different stages of learning? To address this important issue, we analyzed two indicators – the pattern of the success rate of interaction and the average amount of interaction content – during the learning process, see Figs. 9–11.

We plot the patterns of success rates in interaction among the 40 groups in Fig. 9 for system groups and Fig. 10 for random groups. The  $x$ -axis's of Figs. 9 and 10 are groups rank ordered in their performance in the

Table 6  
One-way ANOVA table of interaction success rates

	SS	DF	MS	$F^*$	Significant
SSC	385.52	1	385.52	34.76	.000
SSE	430.13	38	11.32		
Total	815.65	39			

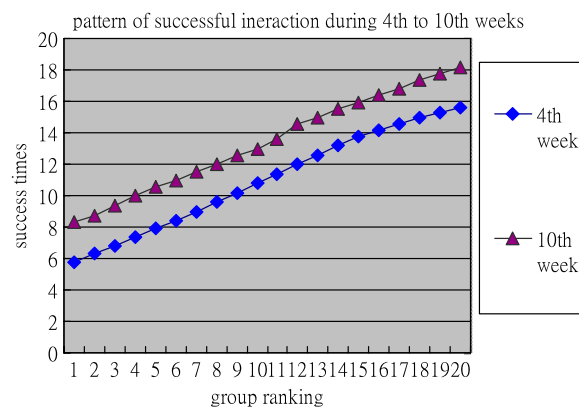


Fig. 9. Systematic-grouping pattern.



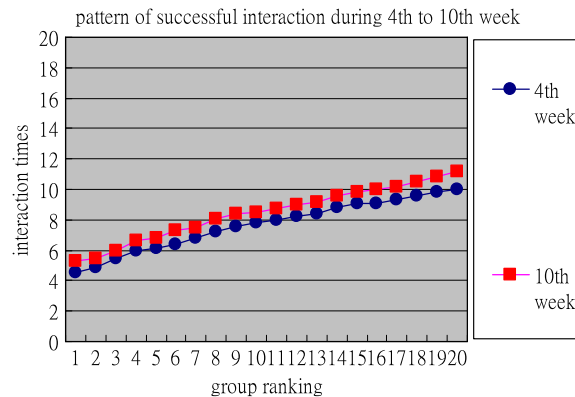


Fig. 10. Random-grouping pattern.

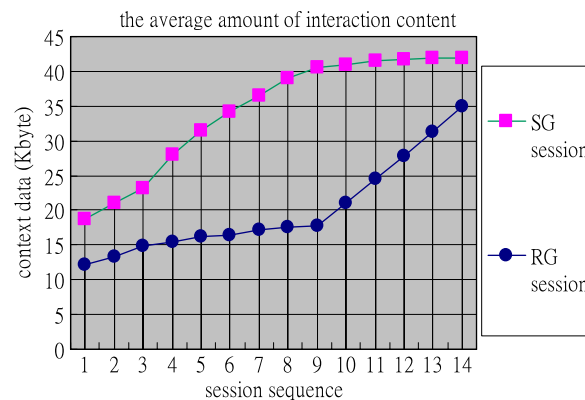


Fig. 11. Pattern of average amount of interaction content.

first stage, with the higher number indicating a higher performance. We continuously observed them from the 4th week to the 10th week to discover any change in their interaction success rate and amount of interaction content. We find that as the 4th week began, the system groups (SG) achieved a higher success rate than the random groups (RG) through the end of the 10th week. Comparing the SG and RG curves, we recognize a distinct difference in growth pattern. The SG members showed on average greater growth (average increase of 2.4 times in 7 weeks) than the RG ones (average increase of 0.9 times in 7 weeks), and at all stages the SG members performed better than the RG ones.

To understand the amount of interaction content between SG and RG students in this period, the content data recorded by our WBI system were extracted for analysis. There were 14 interaction sessions in 7 weeks. We plot pattern of the average amount of interaction content in Fig. 11.

Two observations can be delineated from Fig. 11. First, the average amount of interaction content of system groups is greater than that of the random groups in each session. It seems that students of system groups discuss more and share more of their ideas and information. They exhibit rapid growth of interaction content during the early sessions. Students of random groups do not interact much during the early stages in class, and their growth in interaction content is very slow. The other interesting finding is the interaction tendency in learning. System group students show steady growth in interaction content starting in the 10th session, and then they maintain a stable level until completing the group tasks. To the contrary, the RG students demonstrate little growth in interaction content before the 10th session. Then, suddenly an incredibly high growth of interaction content emerges from the 10th session to the end session. We postulate that RG students procrastinate in completing their group tasks due to low interaction among group members in the early stage of e-learning process. For this reason, they are forced to cram lots of interactions in the last phase.



In terms of group learning and cooperation, field researchers also share similar viewpoints. Group supervision theorists and practitioners stress good group learning results from feedback and discussion with peers and supervisors (Enkenberg, 2001; Rovai, 2001; Wardell & Paschetto, 2001). Thus, instructors of a web-based course should pay more attention to facilitate students in terms of promoting interaction and cooperation, and “warming up” the group climate for learning. (Liaw & Huang, 2000; Powell, Aeby, & Aeby, 2003). In our study, system groups with on-line interaction monitoring are proven more effective than random groups in producing more rewarding and enjoyable learning experience for students on the WBI system.

In sum, our research findings show that both Hypotheses 1 and 2 proposed in the previous section are strongly supported by empirical evidence. That is, groups formed using our grouping algorithm and with interaction monitor are more effective groups and spend more time on discussions in WBI than random groups without interaction monitor.

## 6. Conclusion

Interaction is a critical success factor that affects group learning. It also plays an important role in making students not feel isolated when they learn on the web. There has been little understanding so far in grouping methods and how to effectively promote interaction on the WBI. The need for such understanding becomes more critical now as instructors usually face a heavy enrollment of students in e-learning. In the past, many instructors had to spend much time confronting the difficulties arising from grouping students and promoting interaction among group members in the e-learning setting. Researchers and platform designers of e-learning have to address similar challenge. To solve these practical problems, we propose a useful approach based on data mining techniques to group students and develop an interaction monitoring system to aid instructors in the facilitation of interactive communication. Our novel grouping method based on data mining techniques is shown to produce more effective groups. After the groups have been successfully formed by our method, our on-line interaction monitor captures and analyzes interaction status among group members, providing useful information to the instructors of Web-based courses for more effective facilitation of group interaction in the whole e-learning process.

A field experiment was conducted in a web-based course to validate the effectiveness of our grouping methods and facilitation tools. The groups formed by our method with interaction monitor performed better in all measures such as satisfaction level, system usage, rate of successful interaction, and interaction content. They also exhibited better patterns of interaction. We also received positive response from the instructor, who regarded our grouping method and facilitation tools as appropriate and useful because the instructor needed to devote his attention only to groups and individuals that suffered interaction difficulties. Our on-line interaction monitor provided valuable analytical data to assist the instructor in developing corresponding strategies to facilitate interaction.

In sum, we propose in this study an innovative grouping method and an effective interaction facilitation mechanism for e-learning that are validated by field experiments. The beneficial effects of our approach and the success of our design have been confirmed by both qualitative and statistical analyses. There are several worthy avenues of extensions to our study for future research. For instance, cross-sectional surveys can be applied to test the degree of effectiveness in different curricula among different respondents. Further, it may be possible to test the hybrid modes of different instructors in e-learning. We expect to uncover interesting issues as we continue to probe and explore the potentials of our tools and methods.

## Appendix A

### Satisfaction questionnaire (Gladstein, 1984)

Number	Topics
1	We are satisfied with our group members
2	We are satisfied with the cooperation level among group members
3	Generally speaking, we are all satisfied with this group

## References

- Allport, G. W., Vernon, P. E., & Lindzey, G. (1960). *A study of values*. Boston: Houghton.
- Baer, J. (2003). Grouping and achievement in cooperative learning. *College Teaching*, 51(4), 169–174.
- Blaszczak, W. M. (1990). The place and role of regional culture in social life processes in the upper Silesian community prior to 1945. In J. Szumacher (Ed.), *Upper Silesia: The cultural case in particular* (No. 30). University of Warsaw.
- Brush, T. A. (1998). Embedding cooperative learning into the design of integrated learning systems: rationale and guidelines. *Educational Technology Research and Development*, 46(3), 5–20.
- Daft, R., & Lengel, R. (1986). Organizational requirements, media richness and structural design. *Management Science*, 32, 554–569.
- Dash, M., Liu, H., & Xu, X. (2001).  $1 + 1 > 2$ : merging distance and density based clustering. In *Proceedings of the IEEE seventh international conference on database systems for advanced applications* (pp. 32–39).
- Edwards, J. R. (1996). An examination of competing version of the person–environment fit approach to stress. *Academy of Management Journal*, 39, 292–339.
- Ellis, D. G., & Fisher, A. B. (1994). *Small group decision making*. New York: McGraw Hill.
- Enkenberg, J. (2001). Instructional design and emerging teaching models in higher education. *Computers in Human Behaviour*, 17, 495–506.
- Fulk, J., & Boyd, B. (1991). Emerging theories of communication in organizations. *Journal of Management*, 17, 407–446.
- Fulk, J., Schmitz, J., & Steinfield, C. W. (1990). A social influence model of technology use. In J. Fulk & C. W. Steinfield (Eds.), *Organizations and communication technology*. Newbury Park, CA: Sage.
- Gear, T., Vince, R., Read, M., & Minkes, L. A. (2003). enquiry for collective learning in organizations. *The Journal of Management Development*, 22(1), 88–102.
- Gilbert, J., & Tang, T.-P. (1998). An examination of organizational trust antecedents. *Public Personnel Management*, 27(3), 321–338.
- Gladstein, D. (1984). Groups in context: a model of task effectiveness. *Administrative Science Quarterly*, 29, 499–517.
- Groth, R. (1997). *Data mining: a hands-on approach for business professionals*. NJ: Prentice-Hall PTR, Upper Saddle River.
- Hara, N., Bonk, C. J., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science*, 8, 115–152.
- Holladay, S. J., & Kerns, K. S. (1997). It's a generation thing: exploring intergenerational friendships. Paper presented at the meeting of the National Communication Association, Chicago, IL.
- Johnson, D. W., & Johnson, F. P. (2000). *Joining together: group theory and group skills* (7th ed.). Boston: Allyn & Bacon.
- Jonassen, D. H., & Kwon, H. I. (2001). Communication patterns in computer mediated and face-to-face group problem solving. *Educational Technology Research & Development*, 49, 35–51.
- Kirschner, P. A. (2001). Using integrated electronic environments for collaborative teaching/learning. *Research Dialogue in Learning and Instruction*, 2(1), 1–10.
- Lebie, L., Rhoades, J. A., & McGrath, J. E. (1996). Interaction process in computer-mediated and face-to-face groups. *Computer Supported Cooperative Work*, 4(2–3), 127–152.
- Lee, M. (1993). Gender, group composition and peerinteraction in computer-based cooperative learning. *Journal of Educational Computing Research*, 9(4), 549–577.
- Liaw, S., & Huang, H. (2000). Enhancing interactivity in web-based instruction: a review of the literature. *Educational Technology*, 40(3), 41–45.
- Marwell, G., & Ames, R. E. (1979). Experiments on the provision of public goods, I: resource, interest, group size, and the free-rider problem. *American Journal of Sociology*, 84, 1335–1360.
- Middleton, R. (1976). Regional difference in prejudice. *American Sociological Review*, 41, 94–117.
- Mitchell, S. N., Rosemary, R., Bramwell, F. G., Lilly, F., & Solnosky, A. (2004). Friendship and choosing groupmates: preferences for teacher-selected vs student-selected groupings in high school science classes. *Journal of Instructional Psychology*, 31(1), 20–32.
- Pena-Shaff, J., Martin, W., & Gay, G. (2001). An epistemological framework for analyzing student interactions in computer-mediated communication environments. *Journal of Interactive Learning Research*, 21(1), 41–68.
- Pinquart, M., Wenzel, S., & Sorensen, S. (2000). Changes in attitudes among children and elderly adults in intergenerational group work. *Educational Gerontology*, 26(6), 523–540.
- Pinsonneault, A., & Kraemer, K. (1993). Survey research methodology in management information system: an assessment. *Journal of Management Information Systems*, 10(2), 75–105.
- Polichar, V. E., & Bagwell, C. (2000). Pedagogical principles of learning in the online environment. *Syllabus*, 52–56.
- Powell, J. V., Aeby, V. G., Jr., & Aeby, C. T. (2003). A comparison of student outcomes with and without teacher facilitated computer-based instruction. *Computers & Education*, 40(2), 183–191.
- Quinlan, J. R. (1993). *C4.5: programs for machine learning*. San Mateo: Morgan Kaufman.
- Riel, M., & Harasim, L. (1994). Research perspectives on network learning. *Learning*, 4(2–3), 91–113.
- Rovai, A. P. (2001). Classroom community at a distance: a comparative analysis of two ALN-based university programs. *Internet and Higher Education*, 4, 105–118.
- Scott, D. J., Chanidprapa, S., Seung, W.-Y., Jared, V. B., & Jason, L. F. (2002). Team development and group processes of virtual learning teams. *Computers & Education*, 39(4), 379–393.
- Spranger, E. (1928). *Type of men*. NY: Stechert-Hafner.
- Tuch, S. (1987). Urbanism, region, and tolerance revisited. *American Sociological Review*, 52, 504–510.

- Vrasidas, C., & McIsaac, M. S. (1999). Factors influencing interaction in an online course. *American Journal of Distance Education*, 13(3), 22–36.
- Wanger, J. A. (1995). Studies of individualism-collectivism: effects cooperation in groups. *Academy of Management Journal*, 38, 152–172.
- Wardell, C. S., & Paschetto, G. (2001). Small group instruction in real-time over the web. In *Proceedings of the inter-service industry training simulation and education conference*. Orlando, Florida.
- Webb, N. M. (1982). Group composition, group interaction, and achievement in cooperative small groups. *Journal of Educational Psychology*, 74, 475–484.
- Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. Berliner & R. Calfee (Eds.), *Handbook of research in educational psychology*. London: Prentice-Hall.