

A Systematic Approach for Learner Group Composition Utilizing U-Learning Portfolio

Yueh-Min Huang^{1,2*} and Ting-Ting Wu¹

¹Department of Engineering Science, National Cheng Kung University, Taiwan // ²Department of Applied Geoinformatics, Chia Nan University of Pharmacy and Science, Taiwan // huang@mail.ncku.edu.tw // n9897104@mail.ncku.edu.tw

* Corresponding Author

ABSTRACT

A context-aware ubiquitous learning environment allows applications to acquire diverse learning behaviors of u-learners. These behaviors may usefully enhance learner characteristics analysis which can be utilized to distinguish group learners for further instruction strategy design. It needs a systematical method to analyze u-learner behaviors and utilize learner characteristics for group composition. This paper proposes an effective and systematic learner grouping scheme containing transformation processes from u-portfolios to the proposed Portfolio Grid, creating a learner similarity matrix, and group composition. This study also evaluates intra-group diversity of each resultant heterogeneous group and analyzes learning behavioral patterns acquired from the study experiment. The results indicate that the proposed learner grouping algorithms had positive effects on group composition and interaction between group members for follow-up ubiquitous collaborative learning.

Keywords

Ubiquitous learning, collaborative learning, group composition, learning portfolio

Introduction

During the past several years, educational research has increasingly focused on collaborative learning (CL) pedagogies (Johnson & Johnson, 1999). Many Researches have demonstrated collaborative learning as an effective teaching approach that encourages student learning with high-level cognitive strategies, critical thinking, and positive attitudes (Hsu, 2003; Johnson & Johnson, 1998), benefitting students in terms of achievement, motivation, and social skills (Johnson & Johnson, 1989; Huang, Huang & Fu, 2011).

Collaborative learning integrating computer-based information technology has transformed the learning environment into Computer Supported Collaborative Learning (CSCL) (Inkenet et al., 1999). Rapid development progress of wireless communication and continuing growth of mobile handheld devices has led learning into Mobile Computer Supported Collaborative Learning (MCSCCL) (Danesh et al., 2001;), even Ubiquitous Computer Supported Collaborative Learning (UCSCL) (Hwang, Hsu & Huang, 2007) using the concepts of a novel Ubiquitous Learning (U-learning) environment (Ogata & Yano, 2004). In such a learning environment, it is possible to actively provide ways for identifying right collaborators, right contents and right services in the right place at the right time according to the individual surrounding context information of learners (El-Bishouty, Ogata, & Yano, 2007).

Several studies have indicated that learner group composition has become a fundamental issue in collaborative learning. Many researches (Hooper & Hannafin, 1988; Lin, Huang & Cheng, 2010; Webb, 1982) investigating this subject have shown that different grouping criteria for small groups affects learning performance and social behaviors of grouped members. Heterogeneous group composition not only enhances elaborative thinking, but also leads learners to deeper understanding, better reasoning abilities, and accuracy in long-term retention (Johnson and Johnson, 1999). Webb & Palincsar (1996) further suggested that group composition formed with regard to heterogeneity of members' gender, ability, achievement, social economic status, or race, facilitates heterogeneous group composition in collaborative learning. Since a considerable number of researchers have suggested that heterogeneous grouping promotes positive interdependence, better group performance and effective interaction, this study proposes a heterogeneous grouping method to apply in a ubiquitous collaborative learning environment.

In the current information age, many learning systems provide useful functions for gathering more detailed information about learner situations in learning activities helping teachers obtain a richer understanding of learner behavior (Sakamura & Koshizuka, 2005). The purposeful collection of learning records is called a portfolio, which provides evidence of a learner's knowledge, skills, characteristics and dispositions (Sherry & Bartlett, 2005). Moreover, the learning portfolio supports learning by including an evaluation of collected evidence and reflective

commentaries of prospective learning activities (Struyven, Dochy & Janssens, 2005). Given the rapid development of technology, e-portfolios have been increasingly used as an alternative assessment tool which allows teachers, learners, and parents to understand and evaluate the learning process as well as aiding further learning and growth (Bataineh et al., 2007;). Wang and Turner (2006) proposed that reflection helps students and teachers move beyond seeing the e-portfolio as a mere alternative assessment tool to appreciating its value as a learning strategy. In recent years, the e-portfolio has been used as a diagnostic tool for evaluating and reflecting students' learning during the whole learning process (Adams, Swicegood & Lynch, 2004).

Currently, many ubiquitous learning systems provide functions to collect and record learner behaviors or learning events in learning activities. However, existing grouping methods used in outdoor ubiquitous collaborative learning environments that often utilize gender (Savicki, Kelley & Lingenfelter, 1996;), ability (Saleh, Lazonder & De Jong, 2005), individual psychological features (Tian et al., 2008; Wang, Lin & Sun, 2007), or ethnicity (Cordero, DiTomaso & Farris, 1996) to form learning groups, often ignore learners' various learning behaviors. Therefore, this paper proposes a learner grouping algorithm which utilizes the u-learning portfolio (u-portfolio). Since the u-portfolio contains a variety of learner behavior information, the proposed algorithm obtains a more appropriate learner grouping result that more precisely reflects characteristics of each individual learner. The proposed approach utilizes the concept and technique of Repertory Grid to systematically transform an original u-portfolio into a Portfolio Grid. After constructing a portfolio grid, this work builds a learner similarity matrix. Finally, this work uses the FOCUS concept for the learner grouping algorithm to rapidly distinguish learners into appropriate heterogeneous groups. Instructors can utilize outcomes for further ubiquitous collaborative learning and team working with heterogeneous participant groups.

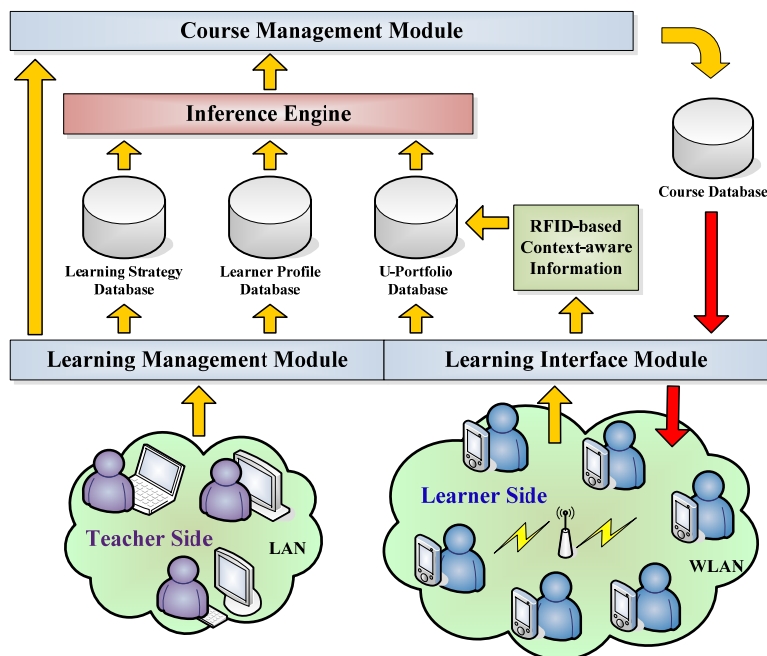


Figure 1. System architecture of the U-plant learning system

U-Plant Learning System

System Architecture and Components

Our previously built u-learning system, U-Plant, collected and recorded the u-portfolio data (Wu, Yang, Hwang & Chu, 2008). Figure 1 shows the U-plant system architecture. The learner interface module provides a friendly and flexible interface for learner operation with mobile devices. The learning processes and learning behaviors are particularly recorded into the u-portfolio database through the learner interface module. The current study also retrieves learning contents and materials from the course database and presents it to learners through the learner interface module. The learning management module enables authorized users, such as the teacher, to configure

learning strategies and create student accounts. Learning strategy configurations such as procedures, rules, and settings are stored in the learning strategy database while the data of created student accounts are stored in the learner profile database. Both learning strategies and student account modifications using the learning management module are updated to the databases. Moreover, the inference engine is responsible for obtaining the required information from databases and making decisions about what learning materials should be retrieved from the course database and delivered to the individual learner. This decision is made by the analysis process in the inference engine with various parameters from the learning strategy, learner profile, and the u-portfolio database. The course management module not only helps the teacher manage learning content, but is also responsible for retrieving adaptive learning content to deliver to students for ubiquitous learning. Context-awareness functionality is a fundamental requirement in the U-Plant learning environment. The learning objects are equipped with tiny RFID tags. Each student takes a mobile device equipped with a passive RFID reader which detects users' location using RFID technology.

Figure 2 shows a learning activity that guided the students to observe the leaf shape of the plant. Learners in the empirical environment of the botanical garden learned, step- by- step, the facts and special characteristics of the empirical objects of learning (plants), according to the learning content and observational steps provided by the system. Learners also used the functions of the system to have interactive discussion and Q&A sessions, record notes, and engage in other learning activities.



Figure 2. Observation interfaces of the U-Plant learning system

U-Portfolio

This study utilizes the web ontology language (OWL) technique for describing learners' u-portfolio and developing a connection between learners and services (Figure 3).

The u-portfolio ontology contains essential learner profiles (Yang, 2006) and the proposed location profile and behavior profile in U-Plant system. Location profile records location movements during students' learning activities utilizing RFID equipments. Behavior profile records u-learning behaviors classified into eight pre-defined categories:

1. Moving: the learner finishes and leaves a certain learning object, and moves on to the next learning object based on the map directions indicated in the mobile device.
2. Losing: the learner cannot successfully locate the appointed learning object to conduct learning.
3. Observing: the learner has moved to a certain learning object and started to use the learning functions in a mobile device until learning is finished and the learner exits.
4. Referencing: the learner looks at the expositions, hints, and notes shown on the learning device.
5. Answering: the learner answers a quiz about observed objects.
6. Interacting: the learner communicates with other classmates.

7. Completing: the learner completes a certain percentage of the learning objects in a specific amount of time.
8. Taking note: the learner utilizes the system support function to record some information during the learning process.

```

Learner ontology = {Profiles, Environment, Devices}
Profiles = {Personal, Calendar, Social, Location, Behavior}
  Personal_profile = {student ID, name, gender, phone, address, email, role}
  Social_profile = {owner, collaborator}
    owner = {student ID, name}
    collaborator = {partner, interactive_type}
      partner = {student ID, name, contact_info}
      interactive_type = {one on one | working_team | community}
  Location_profile = {RFID_ID, site_name, arrival (yyyy:mm:dd;hh:mm), leave(yyyy:mm:dd;hh:mm)}
  Behavior_profile = {moving, losing, referencing, observing, answering, interacting, completing, taking
note}
    moving = {from, reach, time}
      from = {RFID_ID, site_name}
      reach = {RFID_ID, site_name}
      time = {begin(yyyy:mm:dd;hh:mm), end(yyyy:mm:dd;hh:mm)}
    losing = {reach, miss, time}
      reach = {RFID_ID, site_name}
      miss = {RFID_ID, site_name}
      time = {begin(yyyy:mm:dd;hh:mm), end(yyyy:mm:dd;hh:mm)}
    referencing = {object, reference_type, time}
      object = {RFID_ID, site_name}
      reference_type = {exposition | hint | note}
      time = {begin(yyyy:mm:dd;hh:mm), end(yyyy:mm:dd;hh:mm)}
    observing = {RFID_ID, begin(yyyy:mm:dd;hh:mm), end(yyyy:mm:dd;hh:mm)}
    answering = {RFID_ID, correct, begin(yyyy:mm:dd;hh:mm), end(yyyy:mm:dd;hh:mm)}
    interacting = {object, attendee, time}
      object = {RFID_ID, site_name}
      attendee = {student ID, name}
      time = {begin(yyyy:mm:dd;hh:mm), end(yyyy:mm:dd;hh:mm)}
    completing = {ratio, time}
      time = {begin(yyyy:mm:dd;hh:mm), end(yyyy:mm:dd;hh:mm)}
    taking note = {object, behavior, content, time}
      object = {RFID_ID, site_name}
      behavior = {moving, losing, referencing, observing, answering, interacting}
      content = {title, description}
      time = {begin(yyyy:mm:dd;hh:mm), end(yyyy:mm:dd;hh:mm)}
Environment = {School botanical garden, Pool, Library, English classroom, Specialized classrooms}
Devices = {Tool, Equipment}
  Tool = {PDA, mobile phones, Eee PC}
  Equipment = {platform, CPU speed, memory size, screen size}

```

Figure 3. Definition of u-portfolio ontology

Proposed Research Methodology

Portfolio Transformation

Given the flowchart presented in Figure 4, the initial stage retrieves portfolio database records that can be used as factors for distinguishing learners' characteristics. The next step transforms the portfolio data retrieved from LMS into the Portfolio Grid, which column fields consist of a set of *Elements*. Table 1 shows a case of portfolio grid. Each student is listed as an *Element* and put in the top of the grid as a column caption.

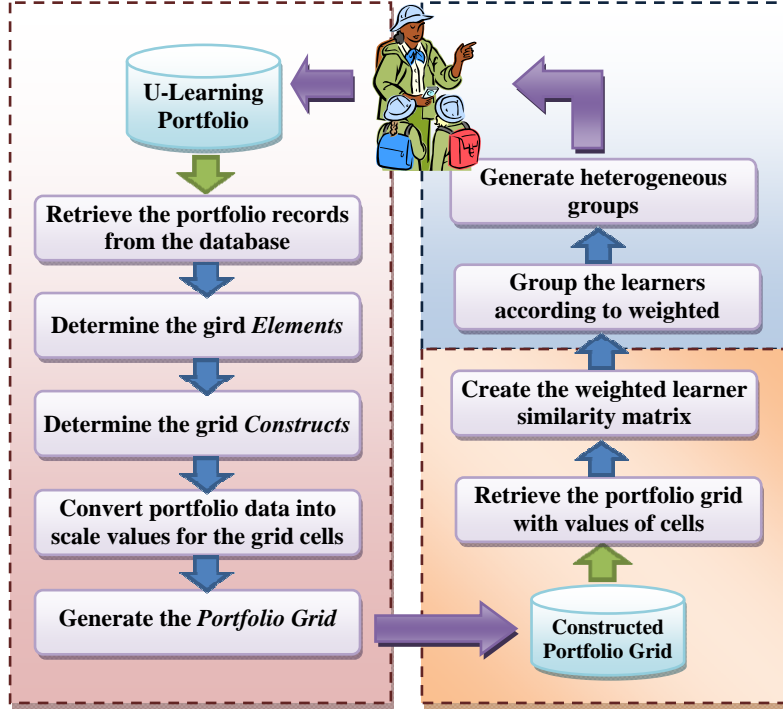


Figure 4. Flowchart of the proposed grouping approach

Table 1. Example of a portfolio grid

<i>Trait Attributes</i>	<i>Elements</i>	<i>Ann</i>	<i>Tom</i>	<i>Jon</i>	<i>Eva</i>	<i>Joy</i>	<i>May</i>	<i>Elements</i>	<i>Opposite Attributes</i>
observing for a long time		5	1	5	5	5	1	observing for a short time	
answering quiz correctly		5	1	1	5	5	5	answer quiz incorrectly	
interacting frequently		5	5	5	1	5	5	interacting seldom	
moving for a long time		1	1	1	5	5	1	moving for a short time	
losing frequently		1	1	5	1	5	5	losing seldom	
answering for a long time		5	1	5	5	3	1	answering for a short time	
referencing for a long time		5	1	1	5	4	1	referencing for a short time	
referencing hint frequently		5	2	1	5	1	1	referencing hint seldom	
referencing note frequently		5	1	3	3	5	2	referencing note seldom	
high degree of completing		1	5	2	1	1	2	low degree of completing	
taking note frequently		5	1	2	4	5	1	taking note seldom	

The rows of a portfolio grid consist of a set of *Constructs* and each *Construct* contains a pair of “trait attributes” and “opposite attributes”. The current work puts trait attributes on the left-hand side of the grid, and opposite ones on the right-hand side. After determining the *Elements* in columns and the *Constructs* in rows, the subsequent procedure provides suitable evaluation values and fills them into the grid cells. The evaluation values are presented using the *K-scale* rating mechanism which converts original data values into *K-scale* rating values using the following calculation formula shown in (1).

$$r = \begin{cases} K, & v = V_{\min} \\ K - \left\lceil \frac{v - V_{\min}}{(V_{\max} - V_{\min} + 1)/K} \right\rceil + 1, & \text{otherwise} \end{cases} \quad (1)$$

where notation r is the obtained *K-scale* rating value by calculation; v is the retrieved average value from a database table field; V_{\min} and V_{\max} are the minimum and maximum average values respectively in the corresponding database table field; and K is the adopted rating scale for the rating calculation. These calculation result values that have lower

ratings indicate a significant characteristic towards the trait attribute. Contrarily, the higher one means a characteristic towards the opposite attribute. In other cases where the retrieved original data is a Boolean-type value, the TRUE and FALSE values are represented using rating values 1 and K respectively.

Create Weighted Similarity Matrix

The subsequent procedure judges the weight of each *Construct* listed in the grid. Different weights mean that each *Construct* can be assigned to a different degree of importance for similarity comparison and heterogeneous grouping. Adjustments to the setting of the weight could be made according to the teachers' experience, or based on the difference between the design of learning activities and group strategy. If the teacher has not set any weights, the system will automatically assign the same weighted value to each construct. Additionally, the settings of the system provide an interface that allows teachers to see a preview of the results of learner groups, which is derived from various weight settings that have been processed by the system's methods of analysis. This provides reference for the teachers in making decisions about setting the weighted value.

Table 2. Example of a similarity matrix

	<i>Ann</i>	<i>Tom</i>	<i>Jon</i>	<i>Eva</i>	<i>Joy</i>	<i>May</i>
<i>Ann</i>	100	30	50	75	66	36
<i>Tom</i>	30	100	57	18	23	70
<i>Jon</i>	50	57	100	39	57	68
<i>Eva</i>	75	18	39	100	59	25
<i>Joy</i>	66	23	57	59	100	52
<i>May</i>	36	70	68	25	52	100

Table 2 shows the matrix of learner similarity generated. Calculation of the weighted value for each element of the similarity matrix in this paper is shown in (2). The notation S_{ij} denotes the percentage of similarity between learners L_i and L_j . The notations n and m are the numbers of *Elements* (learners) and *Constructs* in the portfolio grid. $PG(L_i, C_h)$ denotes the cell value within column (*Element*) L_i and row (*Construct*) C_h in the portfolio grid; α_h denotes the weight assigned to the *Construct* C_h ; and K is the rating scale used in the previous developmental stage of the grid.

$$\forall i, j \in [1, n]$$

$$S_{ij} = 1 - \left(\frac{\sum_{h=1}^m \alpha_h |PG(L_i, C_h) - PG(L_j, C_h)|}{K - 1} / m \right) \times 100 \% \quad (2)$$

, where α_h is the weight of C_h ; $h = 1, 2, 3 \dots m$

Grouping Algorithm

After the weighted similarity matrix of learners is generated, a teacher could determine a threshold of difference or a number of groups for the use of two heterogeneous clustering algorithms in this paper.

Figure 5 shows a heterogeneous grouping with a given difference threshold T . This grouping algorithm not only allows students to cooperate in a more effective manner to complete the assignment, but also enables individual group members to bring distinctive thinking to the design task. The steps of the heterogeneous grouping algorithm are below.

- Step 1. Assign the difference threshold T according to the empirical rule or the learning strategy design.
- Step 2. Due to heterogeneous grouping, first establish a triangular matrix M' as the difference matrix. The element value is 100 minus the element value of the similarity matrix.
- Step 3. Assign every learner to be an independent group by default.
- Step 4. Select the element with the maximum value (the greatest difference value) from all of the elements in the matrix M' .
- Step 5. If the maximum value is greater than the threshold T , find the two learners (row and column of matrix M') corresponding to the element value.

- Step 6. If each of the found learners belongs to a different group, merge them into one group.
- Step 7. Delete the maximum value element just selected. Repeat Steps 4-6 until a difference value greater than T cannot be found in the matrix M' .
- Step 8. Based on the obtained heterogeneous grouping results, the number of people in each group may differ. Therefore, teachers can decide whether or not to set the difference threshold n_T of the number of people. If the difference between any two groups of people is greater than n_T , then teacher can execute the *balance()* function to conduct minor adjustments to balance the number of people.

```

given a similarity matrix  $M = \{ m_{ij} \mid 1 \leq i, j \leq n \}$  where  $n$  is the number of students.
given a difference threshold  $T$  for the heterogeneous grouping
let triangular matrix  $M' = \{ 100 - m_{ij} \mid m_{ij} \in M \text{ and } i < j \}$ 
let  $\{s_1, s_2, s_3, \dots, s_n\}$  denote the set of all students
 $\forall i \in [1, n]$ , let  $C_i = \{s_i\}$ 
let  $C = \{C_i \mid i \in [1, n]\}$ 
while TRUE
    select an element  $m_{ij} \in M'$  such that  $m_{ij}$  is largest in  $M'$ 
    if  $m_{ij} \geq T$ 
        find  $C_h$  and  $C_k$  in  $C$  such that  $s_i \in C_h, s_j \in C_k$ 
         $C_h \leftarrow C_h \cup C_k$ 
        remove element  $C_k$  from  $C$ 
        set  $m_{ij} = \text{NULL}$ 
    else
        exit while
    end if
end while
if  $\exists C_i, C_j \in C$  and  $|C_i| - |C_j| > n_T$  //  $n_T$  indicates the group size error threshold
    call balance( $C, T, n_T$ )
end if
//  $C$  represents the clustering result

```

Figure 5. Heterogeneous grouping algorithm with given difference threshold

```

given a similarity matrix  $M = \{ m_{ij} \mid 1 \leq i, j \leq n \}$  where  $n$  is the number of students.
given an expected number of student groups,  $G$ , for heterogeneous grouping.
let triangular matrix  $M' = \{ 100 - m_{ij} \mid m_{ij} \in M \text{ and } i < j \}$ 
let  $\{s_1, s_2, s_3, \dots, s_n\}$  denote the set of all students
 $\forall i \in [1, n]$ , let  $C_i = \{s_i\}$ 
let  $C = \{C_i \mid i \in [1, n]\}$ 
while  $|C| > G$  and  $\exists m_{ij} \in M'$  such that  $m_{ij} \neq \text{NULL}$ 
    select an element  $m_{ij} \in M'$  such that  $m_{ij}$  is largest in  $M'$ 
    find  $C_h$  and  $C_k$  in  $C$  such that  $s_i \in C_h, s_j \in C_k$ 
    if  $|C_h| + |C_k| \leq (n / G)$ 
         $C_h \leftarrow C_h \cup C_k$ 
        remove element  $C_k$  from  $C$ 
        set  $m_{ij} = \text{NULL}$ 
    end if
end while
if  $\exists C_i \in C$  and  $|C_i| \neq (n / G)$ 
    call balance( $C, n / G$ )
end if
//  $C$  represents the grouping result

```

Figure 6. Heterogeneous grouping algorithm with given number of groups

Figure 6 shows a heterogeneous grouping with a given number of groups G . The steps are as follows:

- Step 1. Assign the expected number of groups, G , according to the empirical rule or the learning strategy design.
- Step 2. Due to heterogeneous grouping, first establish a triangular matrix M' as the difference matrix. The element value is 100 minus the element value of similarity matrix.
- Step 3. Set every learner to be an independent group by default.
- Step 4. If the quantity of all groups is currently larger than G , select an element with the maximum value (the greatest difference) from all of the elements in the matrix M' .
- Step 5. Find the two learners (row and column of matrix M') corresponding to the element value.
- Step 6. If each of the found learners belongs to a different group, merge them into one group.
- Step 7. Delete the maximum value element just selected. Repeat Steps 4-6 until the number of groups equals G .
- Step 8. Based on obtained heterogeneous grouping result, teachers can decide whether or not to set the difference threshold n_T of the number of people and execute the *balance()* function to conduct minor adjustments to balance the number of people.

Simulation and Experiment Evaluations

Simulation Assessment

This section conducts a simulation for evaluating and comparing the average intra-cluster diversity (AID) of clustering results generated by the proposed approach with random clustering, and clustering according to academic achievement. Formula (3) shows the calculation of the value of AID, which is the average of the differences of values of learners in each cluster. G_i represents the set of learners in the first cluster, while G represents the set of all clusters. m_{jk} is the value of similarity between the j th student and the k th student in the learner similarity matrix. Because researchers try to achieve heterogeneous clustering, the goal of the simulation is to compare the difference between the results derived from three types of clustering methods and the learners in the average group. A higher value of AID implies greater heterogeneity, and greater heterogeneity implies better results of heterogeneous clustering.

$$AID = \left(\sum_{i=1}^{|G|} \left(\sum_{j=1}^{|G_i|-1} \sum_{k=j+1}^{|G_i|} 1 - m_{jk} \right) \right) / \left(\frac{|G_i|!}{2!(|G_i|-2)!} \right) / |G| \quad (3)$$

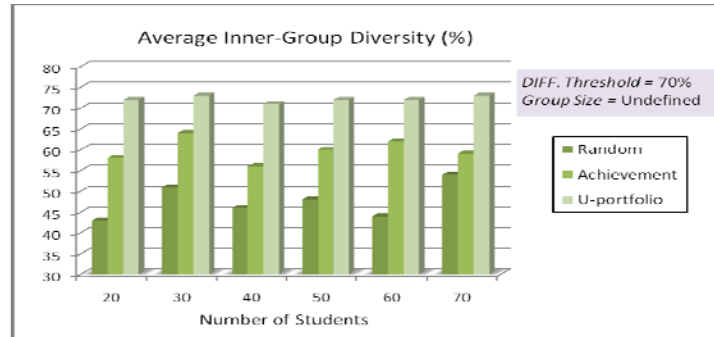


Figure 7. Average inner-group diversity given a difference threshold of 70%

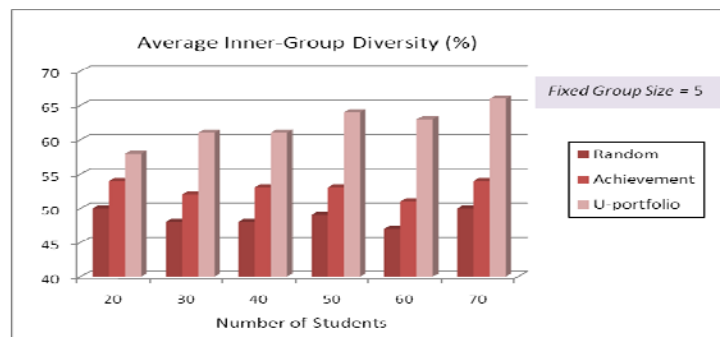


Figure 8. Average inner-group diversity given the respective number of 5

Figure 7 shows the average intra-cluster diversity (AID) of clusters generated with given a threshold of difference parameter of 70 percent. The label 'Random' indicates that random clustering was used to generate groups of learners. 'Achievement' indicates that learners were categorized into groups according to their academic marks. 'U-portfolio' indicates that the systematic method of clustering proposed in this paper, which utilizes the u-learning portfolios, was used. In this case, although the number of students increased, the AID values of the groups resulting from the proposed heterogeneous clustering were greater than the AID values of groups resulting from random clustering and clustering by academic achievement. The AID values of these clusters were also given a threshold of difference at 70 percent. The result of the simulation demonstrates that the proposed heterogeneous clustering with a given difference threshold obtained a higher average intra-cluster diversity compared to random clustering and clustering according to academic achievement. A higher AID value promotes the effect of intra-group learning.

Figure 8 shows the average intra-cluster diversity of generated clusters given the respective numbers of groups. In this simulation, each generated cluster contains five students and the AID values of the proposed method of clustering are greater than AID values resulting from other criteria for clustering (random and academic achievement). The results illustrate that the proposed method generates better heterogeneous clusters for facilitation of collaborative learning activities than the methods of random clustering and clustering according to academic achievement. Results also show that increase in the number of students implies increase in AID value.

Experiment Assessment

Participants

Participants included 114 fourth-graders from four classes in an elementary school in Taiwan. The students consisted of sixty-one males and fifty-three females between the ages of nine and ten years who voluntarily participated in this experiment. These four classes were randomly assigned four different grouping methods. One class was the control group G_R with a random grouping method (27 students). Another class was the control group G_A with an achievement grouping method (30 students), and each of the students was assigned to a heterogeneous group according to their school achievement. The third was the experimental group G_T using the proposed grouping method with a given difference threshold (27 students). The fourth was the experiment group G_N using the proposed grouping method with a given group number (30 students).

Procedures

The school year in Taiwan consists of two semesters. Students in the four classes used the U-Plant learning system for personalized outdoor learning during the first semester and the learning behavior of students was collected and recorded in the u-portfolio database. Because the students had previously used this learning system, researchers did not have to spend a significant amount of time familiarizing students with the operation of the learning system. The experiment was conducted in the second semester and was devoted to lessons on plant biology. The second semester of the school year consists of sixteen weeks minus midterm and final examination weeks. For the first four weeks, students participated in traditional classroom learning with teacher instruction in basic plant biology. In the fifth week, students were given quick training in the operation of the learning system. Students experiencing any operational problems during the training period were given technical assistance by the teacher. At the end of the fifth week, the four classes were randomly assigned the four methods of clustering. From the 6th to 14th week, all grouped students from each class attended three 40-min. lessons per week. During the first two 40-min lessons, each student received a handheld device equipped with Wi-Fi and RFID technology to carry out ubiquitous collaborative learning activities in the school botanical garden. In the last 40-min lesson, students provided feedback about their learning experience and the teacher discussed this feedback with students. Finally, in the 15th and 16th weeks, each group was asked to write a report on learning about plant biology. Each group shared their report with the other groups. To improve the student presentations, these presentations made up 25 percent of the final grade. All groups were asked to upload their final reports and the teacher bulletined the best one. The overall process of experimentation was videotaped to facilitate follow-up investigation and analysis of behavior.

Preliminary of Data Analysis

In our experiment design, four classes were respectively assigned four different grouping methods. The students in control group G_R were randomly grouped into groups of three students. The size of the control group G_A , which was grouped based on school achievement, also assigned three students for each group to carry out outdoor collaborative learning. In experimental group G_T , the teacher assigned a 70 percent difference threshold to divide the class into nine groups with three members in each group. The teacher also assigned an expected grouping number six to the proposed grouping algorithm in which the experimental group G_N generated five students for each group. During the experiment, the collected learning behavior with frequency and sequence data were analyzed using Lag Sequential Analyses (LSA) to find out differences in behavioral patterns according to the four learner grouping methods (Astous & Robillard, 2002; Bakeman & Gottman, 1997).

Experiment Results

This experiment evaluated whether the proposed grouping approach was effective and helpful for learners and whether the heterogeneous grouping method achieved directional and structural behavior patterns to affect traditional collaborative learning behaviors. The learning behavior collected during the experiment activities was classified into eight categories including moving (M), losing (L), referencing (R), observing (O), answering (A), interacting (I), completing (C), and taking note (N). The lag sequential analysis was used to analyze the frequencies and sequences of each student in each group, resulting in four groups shown in Table 3, 4, 5, and 6.

Table 3. Frequency transition table of the control group G_R

	M	L	O	R	A	I	C	T	Total
M	10	55	458	250	0	117	0	97	987
L	52	6	3	187	0	163	0	49	460
O	17	29	42	792	1058	367	332	321	2985
R	193	32	982	129	1177	540	433	556	4042
A	951	5	1169	453	47	388	957	234	4204
I	57	2	273	196	271	56	371	207	1433
C	974	3	61	5	0	129	0	79	1251
T	5	31	78	87	118	77	105	21	522
Total	2259	163	3066	2099	2698	1837	2198	1564	15884

Table 4. Frequency transition table of the control group G_A

	M	L	O	R	A	I	C	T	Total
M	12	46	573	352	0	254	0	105	1342
L	43	7	4	152	0	149	0	37	392
O	21	22	31	985	1123	784	689	357	4012
R	175	19	1065	132	1161	774	646	569	4541
A	1073	5	1226	386	31	412	1083	321	4537
I	63	2	684	217	862	42	389	254	2513
C	1064	5	52	11	0	153	0	89	1374
T	6	23	65	78	125	87	95	17	496
Total	2457	129	3700	2313	3302	2655	2902	1749	19207

Table 5. Frequency transition table of the experimental group G_T

	M	L	O	R	A	I	C	T	Total
M	17	48	628	387	0	683	0	179	1942
L	45	8	32	295	0	487	0	106	973
O	29	13	153	1185	1217	1012	1063	523	5195
R	473	27	1152	673	1279	1026	1173	686	6489
A	1154	6	1342	1053	218	1043	1365	451	6632
I	879	3	1056	986	1197	783	1168	657	6729
C	1379	7	172	159	0	849	0	174	2740
T	11	34	449	298	252	639	209	65	1557
Total	3987	146	4984	4936	4163	6522	4978	2841	32257

Table 6. Frequency transition table of the experimental group G_N

	M	L	O	R	A	I	C	T	Total
M	7	34	472	317	0	774	0	116	1720
L	37	5	29	189	0	389	0	84	733
O	18	6	127	1056	1174	1276	942	462	5061
R	287	21	1103	389	1181	1292	1007	487	5767
A	1035	5	1247	897	196	1213	1156	266	6015
I	997	3	992	739	927	1011	859	318	5846
C	1174	5	78	45	0	1153	0	122	2577
T	8	24	427	258	231	893	199	57	2097
Total	3563	103	4475	3890	3709	8001	4163	1912	29816

Lag sequential analysis indicates the degree of confidence with which one type of data influences the occurrence of another. In other words, LSA observes certain behavior patterns which occur immediately after another behavior occurs. Therefore, this study performed statistics for calculating the sequences of behaviors and frequencies of each sequence of pair behaviors which occurred during the whole learning process. In Tables 3, 4, 5, and 6, the value of each cell with row item $x \in \{M, L, O, R, A, I, C, T\}$ and column item $y \in \{M, L, O, R, A, I, C, T\}$ represents the frequency of occurrence of behavioral pair x and y where behavior y occurred immediately after behavior x . In other words, x belongs to the set of behaviors recorded on the left of the table while y belongs to the set of behaviors recorded at top of the table. After calculating frequency transition tables from collected behaviors data, the subsequence process calculates the statistical significance of observed adjacent behaviors via utilizing the z score proposed by Allison and Liker (1982). The statistic formula of the z score follows:

$$z = \frac{p_{B|A} - p_B}{\sqrt{\frac{p_B(1-p_B)(1-p_A)}{(n-k)p_A}}} \quad (4)$$

where $P_{B|A}$ is the observed proportion of behavior B occurrences at lag k after behavior A occurs; P_A and P_B are the quantity of observed proportions of behavior A and B respectively; n is the sample size of behaviors in the sequence. Tables 7, 8, 9, and 10 show calculation results and present a matrix of z statistics computed for every pair of adjacent behaviors. A calculation result greater than +1.96 indicates a statistically significant continuity level ($p < .05$) with a 95 percent level of confidence. In another words, a significant relationship exists between the two behaviors.

Table 7. Adjusted residuals table (z-scores) of the control group G_R

	M	L	O	R	A	I	C	T
M	-3.72	-1.72	2.23*	0.28	-3.47	0.29	-6.36	-0.16
L	-2.13	-2.42	-4.33	-0.11	-9.72	0.14	-10.40	-3.47
O	-1.97	-1.67	-1.42	8.26*	11.77*	1.42	0.97	0.78
R	0.23	-1.61	11.21*	-0.72	12.77*	4.97*	2.51*	5.13*
A	10.28*	-2.23	13.31*	2.85*	-1.90	1.89	9.87*	0.25
I	-1.39	-3.27	0.37	0.17	0.21	-0.73	1.78	0.19
C	10.97*	-2.97	-3.23	-2.64	-5.17	0.18	-8.26	-0.16
T	-2.13	-5.23	-0.92	-0.12	-0.23	-0.64	-0.19	-1.42

Table 8. Adjusted residuals table (z-scores) of the control group G_A

	M	L	O	R	A	I	C	T
M	-2.16	-1.21	5.23*	1.73	-5.27	0.83	-7.67	0.17
L	-1.37	-2.67	-3.63	0.51	-10.62	0.46	-11.31	-1.64
O	-1.98	-1.85	-1.91	9.76*	12.47*	7.79*	6.85*	1.87
R	0.63	-0.72	11.26*	0.32	12.89*	7.17*	5.97*	4.69*
A	11.64*	-3.13	13.52*	2.16*	-1.76	3.73*	11.93*	1.31
I	-0.69	-4.71	6.42*	0.69	8.87*	-1.48	2.84*	0.97
C	10.97*	-3.47	-0.89	-2.34	-6.46	0.59	-9.23	-0.06
T	-2.84	-2.34	-0.58	-0.37	0.27	-0.18	0.12	-0.94

Table 9. Adjusted residuals table (z-scores) of the experimental group G_T

	M	L	O	R	A	I	C	T
M	-5.83	-0.82	4.21*	1.92	-9.97	5.97*	-10.42	0.72
L	-1.21	-8.19	-1.97	1.12	-13.83	3.23*	-14.72	0.13
O	-3.23	-6.85	0.34	12.26*	12.59*	9.21*	9.97*	3.72*
R	2.92*	-4.34	10.92*	5.34*	12.97*	9.34*	11.97*	6.21*
A	11.28*	-8.97	14.42*	9.64*	0.97	9.42*	14.52*	2.02*
I	7.92*	-9.65	9.72*	8.59*	10.26*	6.85*	11.63*	4.97*
C	13.29*	-8.28	0.48	0.29	-11.28	7.26*	-12.92	0.59
T	-7.62	-2.42	2.31*	1.28	1.03	4.59*	0.92	-0.34

Table 10. Adjusted residuals table (z-scores) of the experimental group G_N

	M	L	O	R	A	I	C	T
M	-6.52	-3.59	2.42*	1.21	-8.83	6.42*	-9.72	-1.05
L	-3.28	-7.15	-3.84	-0.41	-12.97	1.97*	-13.77	-1.59
O	-5.21	-6.88	-0.78	9.42*	11.28*	13.26*	7.97*	2.19*
R	0.85	-4.72	9.84*	1.90	11.85*	13.83*	8.63*	2.82*
A	9.17*	-7.93	12.97*	7.19*	-0.29	12.19*	10.34*	0.73
I	8.28*	-8.64	8.19*	5.90*	7.72*	8.79*	6.72*	1.59
C	10.84*	-7.49	-1.89	-2.92	-10.15	10.15*	-11.40	-0.92
T	-6.26	-4.19	1.92	0.23	0.19	6.89*	-0.19	-2.74

Based on the results of the above z statistics, this study converted the calculations into diagrams of behavioral relationships. Figures 9, 10, 11, and 12 are the behavioral diagrams of the four groups.

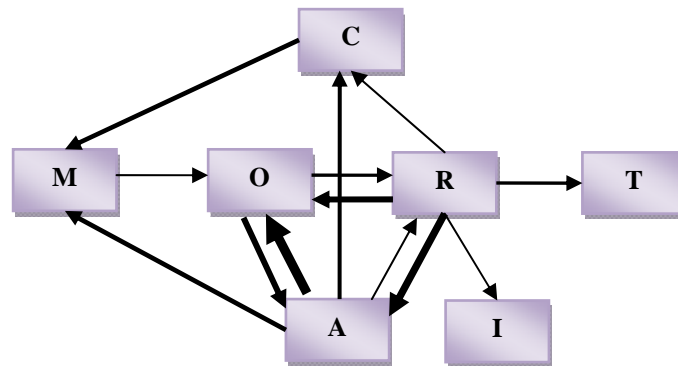


Figure 9. Behavioral transfer diagram of the control group G_R

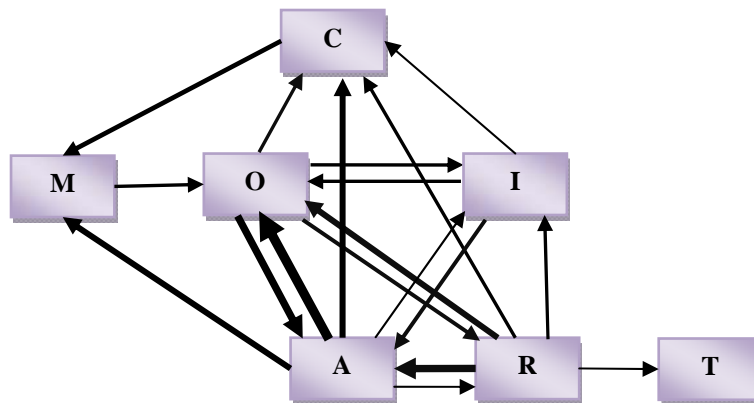


Figure 10. Behavioral transfer diagram of the control group G_A

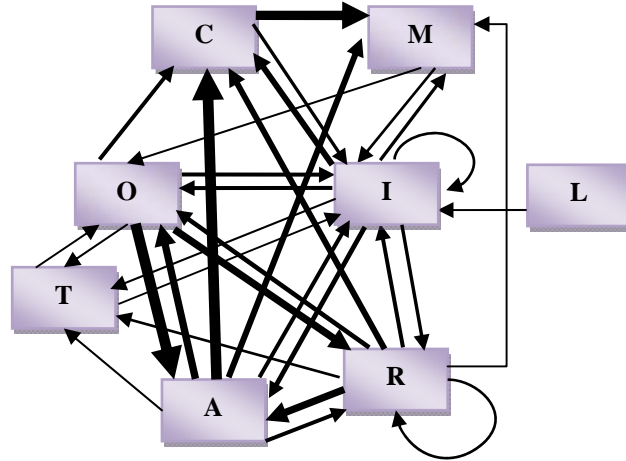


Figure 11. Behavioral transfer diagram of the experimental group G_T

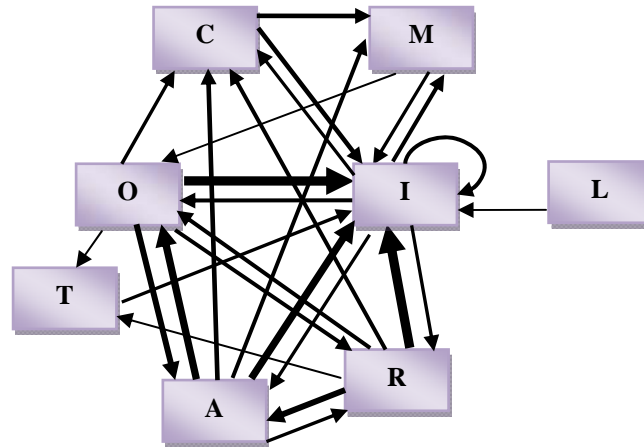


Figure 12. Behavioral transfer diagram of the experimental group G_N

Figure 9 shows a strong link in control group G_R between Answering (A) and Observing (O), indicating that most students observe empirical objects of learning to answer quizzes generated by the learning system. Referencing (R) and Answering (A) are also strongly linked, indicating that students frequently answered the quizzes by referring to the explanations, hints, or notes presented by the learning device. Referencing (R) and Observing (O) also have a strong relationship, indicating that students are more attentive to empirical objects of learning due to the use and assistance of the learning system. In addition, a behavior sequence of $M \rightarrow (O \rightarrow R \rightarrow A)$ is shown in the diagram, and $O \rightarrow R \rightarrow A$ formed a sequence cycle. These results showed that students moving to a correct learning location concentrated on observing the plant with materials or explanations provided by the system, and attempted to finish the tasks.

Compared to the control group G_R , the behavioral pattern of control group G_A (achievement grouping) increased connections and interactions between learning behaviors (Figure 10). Behaviors of Observing (O) and Answering (A) in control group G_R have a stronger relationship with Interacting (I) than do Observing and Answering in control group G_A , indicating that students discuss the learning materials when they observe the plants and answer the questions. The relationship between Answering (A) and Completing (C) in control group G_A is stronger than in control group G_R , showing that the students have better comprehension and degree of completion of assignments when they are grouped using the basic heterogeneous clustering method with student achievement data. This behavioral pattern generates two cycles of $O \rightarrow R \rightarrow A$ and $O \rightarrow I \rightarrow A$, indicating that heterogeneous clustering and clustering according to academic achievement could not only enhance the learning interests of students, but also promote more discussion among students.

Figures 11 and 12 reveal that the behavioral patterns of the experimental groups G_T (grouping with given difference threshold) and G_N (grouping with given number of groups) were more complex than the two control groups, and behavior in the two experimental groups was more brisk and interactive than behavior in the two control groups. Because the learning behavior of learners in the previous u-learning empirical learning environment was considered for heterogeneous clustering, these learners had higher heterogeneity. Therefore, the difference between the learners was also more significant than the difference resulting from clustering in the control group. This significant difference indicated that members of the experimental group had more interactive discussion during the process of group learning and more exchange of opinion and cognitive communication. These results demonstrated that the method of heterogeneous clustering that considered students' behavior in previous U-plant learning activities had more effect on the students' level of cooperation and social interaction in work.

Figure 11 shows that almost all types of behavior were linked with other types of behavior. A possible inference is that many students with higher heterogeneity encourage each other to more actively acquire new skills, ideas and knowledge, and build solutions to educational problems by working together. A strong behavior sequence from $O \rightarrow A \rightarrow C \rightarrow M$ is shown, indicating that the majority of students exhibited greater degrees of accomplishment.

Although Figure 12 resembles Figure 11, comparison of the relationship between behaviors reveals a different intensity of connections between the learning behaviors, particularly connections to the Interacting (I) behavior. The links to Interacting (I) behavior in experimental group G_N were stronger. This phenomenon indicates that much discussion and interaction occurred between the group members, possibly because the students were grouped according to their previous U-plant learning behavior, and the size of each learner group in G_N was larger than the groups in G_R , G_A , and G_T . Due to the larger group size, the degree of mutual interaction required to reach a consensus on learning and seemed more intense and frequent during the learning activity. Although intense discussion could lead to deeper understanding among the students, it could also generate too many distractions that could hinder the accomplishment of tasks and purpose of learning. Therefore, in a u-learning environment with high heterogeneous clustering of learners, the quantity of learning groups and the number of people in each group must still be controlled to avoid an inability to achieve a consensus on learning and reach goals for group learning.

The results above show that the behavioral relationships generated from the two experimental groups are more complex than the behavioral relationships in the two control groups. These results demonstrate that groups with a greater diversity of behavior exhibited more interaction between learners and effected the process of learning more significantly.

Conclusion

A context-aware ubiquitous learning environment has useful functions for gathering data on the learning behavior of students. These u-portfolio data can help enhance the analysis of the behavior, habits, styles, capability, and potential of learners, which could improve learning designs. In addition, research suggests that collaborative learning with heterogeneous group composition positively affects learners in regards to positive interdependence, social skills, interaction, and better group performance.

This paper, proposes a systematic process for analyzing u-portfolios, building portfolio grids, calculating a learner similarity matrix, and generating heterogeneous learner groups for collaborative learning. Moreover, this study attempts to investigate the effects of different methods of group composition that consider the early learning behavior of students, and explores whether the proposed method of clustering significantly influences the behavior of learners in ubiquitous collaborative learning. The first evaluation utilized the simulation technique to analyze the efficacy of intra-cluster diversity with different clustering methods. Results indicated that the AID values of the proposed clustering algorithms were significantly greater than the AID values of clusters produced by other methods of clustering. The second evaluation experimented to evaluate the proposed clustering methods and utilized Lag Sequential Analyses methodology to assess learning behaviors in the experiment. The results indicated that the proposed clustering algorithms generated highly interactive learning behavior. Future research could address further concerns such as utilizing the Fuzzy or Multi-Repertory Grid to improve results and thoroughly experimenting with and analyzing the effects on learning.

Acknowledgments

This work was supported in part by the National Science Council (NSC), Taiwan, ROC, under Grant NSC 97-2511-S-006-001-MY3, and NSC 98-2631-S-024-001.

References

- Adams, P. W., Swicegood, P. R., & Lynch, S. A. (2004). The diagnostician's portfolio: A tool for evaluation and reflection. *Assessment for Effective Intervention, 29*, 53-66.
- Allison, P. D., & Liker, J. K. (1982). Analyzing sequential categorical data on dyadic interaction: A comment on Gottman. *Psychological Bulletin, 97*, 393-403.
- Astous, P. d', & Robillard, P. N. (2002). Empirical study of exchange patterns during software peer review meetings. *Information and Software Technology, 44*, 639-648.
- Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis* (2nd Ed.), Cambridge, UK: Cambridge University Press.
- Bataineh, R. F., Al-Karasneh, S. M., Al-Barakat, A. A., & Bataineh, R. F. (2007). Jordanian preservice teachers' perceptions of the portfolio as a reflective learning tool. *Asia-Pacific Journal of Teacher Education, 35*, 435-454.
- Blatchford, P., Baines, E., Kutnick, P., & Martin, C. (2001). Classroom contexts: connections between class size and within class grouping. *British Journal of Educational Psychology, 71*(2), 283-302.
- Cohen, E. G. (1994). Restructuring the classroom: conditions for productive small groups. *Review of Educational Research, 64*(1), 1-35.
- Cordero, R., DiTomaso, N., & Farris, G. F. (1996). Gender and race/ethnic composition of technical work group: relationship to creative productivity and morale. *Journal of Engineering and Technology Management, 13*, 205-221.
- Danesh, A., Inkpen, K.M., Lau, F., Shu, K., & Booth, K.S. (2001). *GeneyTM: designing a collaborative activity for the palm handheld computer*. In Proceedings of the Conference on Human Factors in Computing Systems (CHI 2001), Seattle, USA (pp. 388-395).
- El-Bishouty, M. M., Ogata, H., & Yano, Y. (2007). PERKAM: Personalized Knowledge Awareness Map for Computer Supported Ubiquitous Learning. *Educational Technology & Society, 10*(3), 122-134.
- Genesereth, M. (1997). An agent-based framework for interoperability. In J. M. Bradshaw (Ed.), *Software agents* (pp. 317-345). Menlo Park, Calif.: AAAI Press.
- Hooper, S., & Hannafin M. J. (1988). Cooperative CBI: The effects of heterogeneous versus homogeneous group on the learning of progressively complex concepts. *Journal of Educational Computing Research, 4*(4), 413-424.
- Hsu, Y. C. (2003). *Evaluation theory in problem-based learning approach*. Papers Presented at the National Convention of the Association for Educational Communications and Technology, Houston, TX.
- Huang, T. C., Huang, Y. M., & Fu, F.Y. (2011). Cooperative weblog learning in higher education: its facilitating effects on social interaction, time lag, and cognitive load. *Educational Technology & Society, 14*(1), 95-106
- Hwang, W. Y., Hsu, J. L., & Huang, H. J. (2007). *A Study on Ubiquitous Computer Supported Collaborative Learning with Hybrid Mobile Discussion Forum*, Mlearn2007, Melbourne, Australia.
- Inken K. M., Ho-Ching W., Kuederle O., Scott S.D. & Shoemaker G. (1999). *This is fun! We're all best friends and we're all playing: supporting children's synchronous collaboration*. Proceedings of Computer Supported Collaborative Learning (CSCL) '99, Stanford, CA, December.
- Johnson, D. W., & Johnson, R. T. (1998). Cooperative learning and social; interdependence theory. In R. S. Tindale et al. (Eds.), *Theory and research on small groups* (pp. 9-35). New York: Plenum Press.
- Johnson, D. W., & Johnson, R. T. (1999). *Learning together and along: Cooperative, competitive, and individualistic learning* (5th ed.). Boston, MA: Allyn & Bacon.
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computersupported collaborative learning environments: a review of the research. *Computers in Human Behavior, 19*, 335-353.
- Lin, Y. T., Huang, Y. M., & Cheng, S.C. (2010). An automatic group composition system for composing collaborative learning groups using enhanced particle swarm optimization. *Computers & Education, 55*(4), 1483-1493.

- Ogata, H., & Yano, Y. (2004). *Knowledge Awareness Map for Computer-Supported Ubiquitous Language-Learning*. Proceedings of the The 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'04).
- Sakamura, K., & Koshizuka, N. (2005). *Ubiquitous Computing Technologies for Ubiquitous Learning*. Proceedings of the 2005 IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'05), Tokushima, Japan.
- Saleh, M., Lazonder, A. W., & De Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement and motivation. *Instructional Science*, 33, 105–119.
- Savicki, V., Kelley, M., & Lingenfelter, D. (1996). Gender and group composition in small task groups using computer-mediated communication. *Computers in Human Behavior*, 12(2), 209–224.
- Sherry, A. C., & Bartlett, A. (2005). Worth of electronic portfolios to education majors: A “two by four” perspective. *Journal of Educational Technology Systems*, 33, 399–419.
- Silverman, B. (1995). Computer Supported Collaborative Learning (CSCL). *Computers & Education*, 25(3), 81-91.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education*, 30(4), 325–341.
- Tian, F., Wang, S., Zheng, C. & Zheng, Q. (2008). *Research on E-learner Personality Grouping Based on Fuzzy Clustering Analysis*. Proceedings of the 12th International Conference on Computer Supported Cooperative Work in Design.
- van Aalst, J., & Chan, C. K. K. (2007). Student-directed assessment of knowledge building using electronic portfolios. *Journal of the Learning Sciences*, 16, 175-220.
- Wang, D. Y., Lin, S. J., & Sun, C. T. (2007). DIANA: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups. *Computers in Human Behavior*, 23, 1997-2010.
- Wang, S., & Turner, S. (2006). Learning experiences in developing electronic portfolios. *International Journal of Information and Communication Technology Education*, 2, 75-86.
- Webb, N. M. (1982). Student Interaction and learning in small groups. *Review of Educational Research*, 52(3), 421-455.
- Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. Berliner & R. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 841–873.), New York: Macmillan.
- Wu T. T., Yang T. C., Hwang G. J., and Chu H. C. (2008). *Conducting Situated Learning in a Context-Aware Ubiquitous Learning Environment*. Proceedings of the 5th International Conference on Wireless, Mobile and Ubiquitous Technologies in Education (WMUTE2008), Beijing, China.
- Yang, S. J. H. (2006). Context Aware Ubiquitous Learning Environments for Peer-to-Peer Collaborative Learning. *Journal of Educational Technology and Society*, 9(1), 188-201.