Different Identity Revelation Modes in an Online Peer-Assessment Learning Environment: 
Effects on Perceptions toward Assessors, Classroom Climate and Learning Activities

Fu-Yun Yu¹ & Chun-Ping Wu²

¹ Corresponding author
Institute of Education
National Cheng Kung University
No. 1, University Road, Tainan City 70101, Taiwan
Phone: +886-6-275-7575 ext. 56225;
Cell phone: +886-930960052
Fax: +886-6-276-6493
E-mail: fuyun.ncku@gmail.com

² Department of Educational Technology
TamKang University
151 Ying-chuan Road, Tamsui, Taipei County, Taiwan
E-mail: cpwu303@gmail.com

Abstract

The effects of four different identity revelation modes (three fixed modes: real name, anonymity, nickname and one dynamic user self-choice mode) on participants’ perceptions toward their assessors, classroom climate, and past experience with the learning activity in which they were engaged were examined. A pretest-posttest quasi-experimental research design was adopted. Eight fifth-grade classes (age 10-11, N=243) were randomly assigned to four different identity revelation modes in order for them to participate in the study. An online learning system that allows students to contribute to and benefit from the process of question-generation and peer-assessment was adopted. Data analysis confirmed that different identity modes lead participants to view their assessors differently. Specifically, participants assigned to the self-choice and real-name identity revelation modes tended to view their assessors more
favorably than those in the anonymity and nickname groups. The empirical significance of the study as well as suggestions for learning system development, instructional implementation and future study are provided.

**Keywords:** applications in subject areas; computer-mediated communication; cooperative/collaborative learning; human-computer interface; teaching/learning strategies
Different Identity Revelation Modes in an Online Peer-Assessment Learning Environment: Effects on Perceptions toward Assessors, Classroom Climate and Learning Activities

1. Introduction

Empirical evidence generally supports peer-assessment for the promotion of higher-order thinking and cognitive re-structuring in learners (Topping, 1998; Topping, 2009). Nevertheless, a sense of insecurity, heightened anxiety and emotional tension that might affect learners’ perception of and devotion to a learning activity has also been reported (Orsmond, Merry, & Reiling, 1996). Whether such negative emotions associated with a peer-assessment activity can be mitigated by different identity revelation modes is the focus of this study. Considering the fact that with the support of networking technologies, the interacting party’s identity can be easily concealed and/or manipulated dynamically, the effects of different identity revelation modes (real-name, nickname, anonymity and self-choice) on learners’ perceptions toward their assessors, classroom climate, and the learning activity were investigated.

1.1 Theoretical and Empirical Bases of Peer-Assessment and Its Current Research States

Peer-assessment has gained increasing research attention since the 1990s as an alternative pedagogical and assessment model for the promotion of learning (Boud, Cohen, & Sampson, 1999; Hargreaves, 1997). Peer-assessment activities stimulate students to initiate critical thinking to not only provide objective judgments of the quality of the work being evaluated, but also to render constructive comments that may assist in the identification of knowledge gaps and the
improvements to the examined work. In addition, to enable comments to be accepted by peers, both social and argumentation skills (e.g., persuasion) as well as substantial knowledge in the subject area being evaluated are essential (Liu et al., 2001; van Gennip, Segers, & Tillema, 2009). While observing peers’ work, students are alerted to re-examine their own work and make necessary modifications. On the other hand, when students receive feedback from assessors, the comments provided may introduce cognitive conflict and may also direct students to confront their existing cognitive defects. Knowledge structuring and re-structuring are cultivated through the cognitive and discursive processes involving deeper elaboration of materials, self-reflection, comparison, clarification, adjustment, and so on. The aforementioned processes, based on cognitive conflict theory, social constructivism, social learning theory, information processing theory, and metacognition, should promote the development of critical thinking, knowledge integration, and cognitive and metacognitive ability (Falchikov & Goldfinch, 2000; Topping, 1998; van Gennip, Segers, & Tillema, 2009).

In addition to cognitive gains, peer-assessment holds affective benefits for students. The fact that someone other than teachers will view their work tends to induce students to invest more in the preparation of their initial work and the subsequent revision of the work, which is a manifestation of increased motivation (Humphrey, Greenan, & McIlveen, 1997; Macpherson, 1999). Empirical evidence, spanning more than two decades, further substantiates the facilitative effects of peer-assessment on learner motivation, as well as on their sense of responsibility, higher-order thinking skills, cognitive re-structuring, level of performance and attitudes (Brindley & Scoffield 1998; Falchikov & Goldfinch, 2000; Gatfield, 1999; Hanrahan & Isaacs, 2001; Purchase, 2000; Topping, 1998; Tsai, Lin, & Yuan, 2002; van Gennip, Segers, & Tillema, 2009).
While theoretical and empirical foundations are generally supportive of peer-assessment, (van Zundert, Sluijsmans, & van Merriënboer, 2010; Venables & Summit, 2003; Wen & Tsai, 2006; Wen, Tsai, & Chang, 2006; Topping, 1998; Topping, 2009), the importance of clear and pre-specified criteria for peer-assessment that ensure objective assessment of student performance and the provision of adequate feedback leading to further improvement of submitted work has frequently been stressed (Crane & Winterbottom, 2008; Gielen et al., 2010, Orsmond, Merry, & Reiling, 1996). Traditionally, peer-assessment is conducted in a face-to-face or paper-and-pencil format. Fairness is frequently questioned by participants. Also, heightened anxiety and emotional tension associated with the activity have been reported (Orsmond, Merry, & Reiling, 1996; Purchase, 2000; Searby & Ewers, 1997). Additionally, the collection and compilation of ratings and comments associated with such activities are especially tedious, cumbersome and time-consuming, resulting in resistance to swift adoption on the part of teachers.

The many distinct features and capabilities of networked technologies intended to maximize the learning effects of peer-assessment enable assessors to provide timely and constructive feedback about individual student’s work. Features such as high processing speed, immense storage space, learner control, multimedia, simultaneity, instantaneity, space-, time- and device-independence, flexibility, interactivity, and so on are being taken advantage of most frequently (Bonham et al., 2000; McDonald, 2002; Smaldino, Lowther, & Russell, 2008). For the past decade, numerous online peer-assessment systems (for instance, CAP, NetPeas, Vee heuristic, Web-SPA, SWoRD) have been developed and have demonstrated their efficacy with regard to carrying out associated tasks efficiently, effectively and affectively (Cho & Schunn, 2007; Sung et al., 2005; Tsai et al., 2001). Distinctively, with the rapid development of computer
and network technologies and pedagogical features (like online submission, time-stamping, automatic assignments, individual and multiple peer reviews, asynchronous and synchronous interaction, anonymity, learner portfolios, automatic notification, multidimensional evaluations with specific pre-set criteria, multiple types of feedback, and reciprocal peer reviews, among other attributes), the identified drawbacks associated with peer-assessment have been greatly mitigated (Cho & Schunn, 2007, Davies, 2000, Lin et al., 1999; Lu & Bol, 2007; Sung et al., 2005; Trautmann, 2009; Tsai et al., 2001; Yu, 2009).

While great strides have been made in the area, a method by which to moderate the negative emotional reactions associated with peer-assessment is still a topic in need of investigation. In light of the fact that identity disclosure and concealment can be easily managed in online interaction spaces, the comparative effects of different identity revelation modes on the peer-assessment experience is hence the focus of this study.

1.2 Why Different Identity Revelation Modes?

Literature on social psychology suggests that identified and anonymous interactions differ in many psychologically significant ways. Specifically, social scrutiny, evaluation anxiety and self-validation may become less of a concern and threat to participating parties in unidentified situations (Cooper et al., 1998; Franzoi, 2006; Lea, Spears, & de Groot, 2001; Moral-Toranzo, Canto-Ortiz, & Gómez-Jacinto, 2007; Palme & Berglund, 2002) and might theoretically influence learners’ sense of psychological safety (Yu & Liu, 2009). Psychological safety, a shared belief that it is safe to take an interpersonal risk, influences learning from peer-assessment (van Gennip, Segers, & Tillema, 2010). Specifically, high psychological safety encourages
learners to perceive differences in viewpoints as opportunities and thus enhances the quality of learning outcomes. The study of van Gennip, Segers and Tillema (2010) confirmed that when compared with students in a teacher assessment setting, learners in a peer-assessment setting perceive higher psychological safety, thus in turn making it possible to predict their perceptions related to the activity they are engaged in as well as their learning.

On the other hand, according to deindividuation theory, anonymity is conducive to creating a state where individuals submerged in the group often do not think of others as individuals and, similarly, they themselves do not feel that they can be singled out by others (Diener et al., 1976; Festinger, 1950; Gergen, Gergen, & Barton, 1973; Jessup & Tansik, 1991). As a consequence of diminished awareness of self, reduced internal moral restraints and a decreased sense of individual accountability, socially undesirable behaviors, including emotional, impulsive, irrational, regressive and intense disinhibited behavior, are more frequently observed in such loss of identity situations (DeSanctis & Gallupe, 1987; Kiesler, Siegel, & McGuire, 1984).

Furthermore, as gleaned from the literature on media effects and human factors in telecommunications, it is possible to transmit different types of messages of varying lengths in face-to-face and mediated interaction situations. Specifically, in addition to relying on a written message conveyed in an online space, interacting parties in identifiable face-to-face situations, as compared to un-identifiable ones, can resort to additional verbal, non-verbal and/or contextual cues that transpire during the process (Williams, 1977), which may affect the interaction process as well as their views toward the experience.

As has been explained, there seems to be a basic theoretical disagreement as to how readily identifiable (real name), less identifiable (nickname) and un-identifiable (anonymity) real-time
interaction situations may produce differential effects on participants’ perceptions towards their interaction counterparts, the interaction space and the experience. On one side of the argument are researchers who stress the importance of psychological safety and the prospect of minimizing evaluation apprehension and lessening diversion conflict, which is usually hard to attain in identifiable situations. From this perspective, it is posited that being anonymous while participating in a learning activity will allow individuals to focus more on the task at hand (in peer-assessment situations, to provide constructive feedback), rather than being bombarded with emotional burdens that concern what and how peers might think of them in response to their performance on the activity (Pepitone, 1980; Williams, 1977). On the other side of the argument, from deindividuation theorists, it is proposed that anonymity may lead to socially undesirable behavior due to a loss of the sense of self-awareness and individual accountability. From this perspective, un- and less-identifiable interaction situations may intensify anti-normative and disinhibited behavior, which may result in less favorable perceptions towards the interaction parties, the associated activity and the atmosphere in which the interaction takes place.

1.3 Focus and Research Questions of the study

To date, hardly any attention has been paid to the role and effects of identity revelation modes in peer-assessment studies. Studies have found participants exhibit statistically differential preferences toward real-name, anonymity and nickname modes in online peer-assessment situations (Yu & Liu, 2009). However, comparative effects have rarely been under stringent investigation. Furthermore, dynamic identity displays that are under user control seem conceptually intuitive, and they are easily accomplishable technologically. Nonetheless, few, if
any, existing online learning systems have adopted a design incorporating multiple and dynamic identity displays that are under user control. Humanism recognizes the fact that each student is unique, having different motivation, interests, needs, and so on (Ediger, 2006); therefore, schools are recommended to provide adaptive instructions in order to cater to the differences in individual students. While accommodating learner needs (in this case, their need for being or not being identified in collaborative learning situations) is in alignment with the ideology of humanism, the effects, as compared to a fixed identity mode, are not yet known and thus are worth examination.

Hence, in this study, an effort was made to examine the effects of four different identity revelation modes (three fixed modes: real name, anonymity, nickname and one dynamic self-choice mode) on students’ perceptions towards their assessors, their classroom climate, and the learning activity in which they were engaged within an online peer-assessment learning environment. Three research hypotheses are therefore proposed:

1. There will be significant differences in student perceptions toward their assessors among students assigned to four different identity revelation modes.

2. There will be significant differences in student perceptions toward classroom climate among different treatment conditions.

3. There will be significant differences in student attitudes toward past experience with a learning activity among different treatment conditions.

A couple of questions can be investigated along this line. For instance, will socially undesirable behaviors, as suggested by deindividuation theorists, manifest themselves within
un-identifiable (anonymity) or less-identifiable situations (nickname) to an extent that leads to less favorable perceptions of the interacting parties and the associated activity than will occur in identifying arrangements (real-name)? Or, on the contrary, will the heightened psychological safety permissible in anonymous situations result in better perceptions toward the interacting parties and past learning experience as compared to identifiable situations? Knowing that in identifiable situations, interaction is not limited and constrained to an online space, will this extra out-of-online space interaction reach a degree where classroom climate (e.g., the orderliness of the class, the management of the task and learning time, and so on) is compromised? Finally, will accommodating user preferences for different identity revelation modes by allowing students to choose any specific identity revelation mode they wish result in participants perceiving the interaction parties, the learning environment and the engaged activity more favorably as compared to a fixed identity mode that is pre-set by the instructor? The findings from this research should have empirical, theoretical, pedagogical and online system development significance.

2. Methods

2.1 The Learning System

Several online systems that allow students to contribute to and benefit from the process of question-generation and peer-assessment are in existence (for instance, QSIA by Barak & Rafaeli, 2004; MCIDA by Fellenz, 2004; ExamNet by Wilson, 2004; QPPA by Yu, Liu & Chan, 2005; PeerWise system by Denny et al., 2008; Hot Potatoes from Half-Baked Software, 2009). Most systems allow students to generate different types of questions using different media formats,
and some have embedded the peer-assessment and anonymity features. For the purpose of this study, a learning system, Question Authoring and Reasoning Knowledge System (QuARKS) (Yu, 2009), which allows different user identification revelations (particularly, real name, nickname and anonymity) and self-choice selection modes under different learning situations, was adopted. This feature of multiple and dynamic identity revelation modes, the focal point of the investigation, to the best knowledge of the authors, is not yet available in other existing systems. A brief description regarding the conceptual framework and design principles of QuARKS is presented as follows:

QuARKS accentuates technological scaffolding to support student question-generation activities in a timely, flexible and logistically feasible fashion. The conceptual framework proposed by Lin et al. (1999) that is composed of four main dimensions—reflective social discourse, process prompts, process displays and process models—was adopted for the design and development of the various design features of QuARKS. Additionally, in order to support instructors in adjusting different dimensions of the available functions to suit individual instructors’ educational goals and instructional plans in different situations, customizability was highlighted. Explicitly, QuARKS permits instructors to dynamically adjust the types of question-generation activities students are engaged in, the kinds of sub-systems/functions students have access to (i.e., question-generation, question-assessment, question-viewing, drill-and-practice), the specific set of criteria that are determined to be suited for the specific target users for the question-assessment activity, the specific set of generic question stems and accompanying sample questions for reference for question-generation, the specific interaction mode within which question-author and assessor communicate (i.e., assessor-to-author one-way mode, assessor-and-author two-way mode and among various assessors-and-author multi-way
mode), and specific user identification revelation modes (fixed or self-choice), among other options (Yu, 2009).

For the purpose of the study, with the exception that different identity revelation modes were set for different classes assigned to different treatment conditions, accessible functions and guidelines for the activity were kept identical for all treatment groups. Explicitly, students in this study had access to question-generation and question-assessment sub-systems. Considering that true/false and multiple-choice questions are among the most frequently encountered question types in primary schools in Taiwan, these two types of question-generation options were open for all treatment groups. Furthermore, because of a previous study that is supportive of the multi-way interaction mode (Yu, 2011), this option was set for all groups.

Essentially, to enhance interaction, collaboration and negotiation of meaning among the author and all assessors assessing the same question, QuARKS allows the set of interacting parties to communicate back and forth electronically and simultaneously about the appropriateness of an examined question item and to provide subsequent comments about this item in a cyclical form. Assessors are able to give their evaluative feedback using an online assessment form (with assessment criteria and a type-in open space) (see Fig. 1). Alternatively, question-authors can provide an elaborated-upon explanation to which their assessors can further respond (see Fig. 2). Again, in this study, the set of criteria in both assessment forms was the same for all treatment groups.

Finally, to make the collaborative process as fluid as possible, a notification system was put in place. Authors and assessors were alerted automatically by a blinking red icon on their screen when the status of the question in the examination changed. By clicking on the blinking icon and
the built-in “proceed” button, students were transferred directly to the newly posed message where they could take appropriate actions (e.g., the author revising the question based on the assessor’s suggestions and giving evaluative feedback on the performance of the assessor, the assessor re-assessing the quality of the revised question and providing additional comments, if deemed appropriate and useful, etc.). Through an engagement in a cyclical argumentation process with peers, in light of cognitive conflict theory, social constructivism, social learning theory, information processing theory, and metacognition, it was expected that learner understanding of the materials would be enhanced.

![Peer Assessment Form](image)

Only available in the self-choice identity group.

Figure 1. Assessor-to-author assessment form (for assessors to provide feedback to question-authors)
Figure 2. Author-to-assessor assessment form (for question-authors to communicate with assessors)

2.2 Research Design

A pretest-posttest quasi-experimental research design was adopted. Four treatment conditions were devised for the purposes of the study. In the “real-name” condition, when questions and feedback were viewed, the student’s full-name was retrieved automatically from the database and was shown in the “Question-Author” field at the top of the questions and in the “Assessor” field at the top of each comment, respectively. In the “anonymity” group, information on the question-author or assessor was not shown, and only the word “anonymous” was marked in the respective field. In the “nickname” group, the student’s self-created identity was shown instead. Note that users were free to change their nicknames each time they generated or assessed a new question. Finally, to take into account recent findings indicating that participants exhibit statistically significantly different preferences for different user identity revelation modes when providing comments (Yu & Liu, 2009), a “self-choice” group was created. In the self-choice group, rather than being assigned a specific and fixed identity revelation mode, the kind of
information on identity that was displayed was contingent upon the choice of the individual user (See Figure 3). Each time users generated or assessed a new question, they were allowed to select a different identity presentation mode before proceeding further (see Figs. 1 & 2 top portion: identity revelation mode). Eight participating classes were randomly assigned to four different treatment conditions (i.e., two classes per group).
Figure 3: Interaction within the self-choice mode
2.3 Experimental Procedure

Two hundred and forty-three 5th graders (age 10-11) from eight intact classes participated in the study for six consecutive weeks. Eight classes were randomly assigned to four different treatment conditions. To ensure that participants possessed the fundamental skills necessary for the activity to be undertaken, a training session on generating the selected question types and a follow-up peer-assessment with a hands-on activity was arranged at the commencement of the study.

For the duration of the study, students headed to a computer laboratory each week to participate in a 40-minute learning activity after attending three instructional sessions allocated for science. Two units were covered. They were: “Heat Transfer: Conduction, Convection, Radiation and Insulation” and “Sound and Musical Instruments.” For the online learning activity, every week, students were directed first by the instructor to individually compose at least one question for each of the two question types in accordance with the instructional content covered that week. The “main ideas” approach proposed by Dreher and Gambrel (1985) and Ritchie (1985) was introduced. The procedural steps involved were: First, students were directed to identify the main idea of the studied content together with significant details; second, they were asked to form questions that asked for new examples of the identified idea or to write a question about a concept in a paraphrased form if finding new examples for the identified idea proved to be difficult.

After the question-generation activity, students moved on to individually assess at least two questions from a pool of peer-generated questions for each question type. Students were advised
to do the following: First, they were asked to assess the overall quality of the generated question on a five-point rating scale (from “well below average” to “well above average”) and to rate their recommendation of the question to be included in the drill-and-practice item bank (“don’t recommend,” “recommend with reservations” or “highly recommend”); second, to provide focused, objective and constructive feedback by referring to a set of built-in criteria (for instance, unclear question-stems, overly-complicated question-stems, excessively verbose options, multiple correct answers, elusive phrasing, implausible distracters, concise question-stem and option, important concepts, and so on) and to type in any suggestions for the question-author to consider in the designated feedback space (See Fig. 1). Once the assessor-to-author form was sent, the notification system automatically alerted the question-author in such a way that the assessor’s feedback could be instantly retrieved, and responses to received feedback could also be addressed using the author-to-assessor form (see Fig. 2).

To establish a baseline on student perceptions towards their assessors, the classroom climate, and the activity, the real-identity mode was set for all treatment conditions for the first two sessions. A questionnaire on these variables was disseminated for individual completion before different treatment conditions were implemented in different groups at the beginning of the third week. After exposure to the activity for six weeks, all students completed the same questionnaire.

2.4 Measures

The effects of different treatment conditions on students’ perceptions toward their assessors, the classroom climate, and the learning activity were assessed by the same pre- and post-questionnaire, which consisted of three 5-point Likert scales (5=strongly agree, 4=agree,
“Perception toward assessors” (18 items) was defined as the students’ interpretation of their relationship with the assessors, the personality attributes of the assessors, and the competency of their peers as assessors. “Classroom climate” (7-item) was defined as the students’ observations of the overall atmosphere of the class (as being in good order, well-regulated, noisy, students being mostly focused on the task and well behaved, the time spent in class being well managed, and so on), while “Perceptions toward the engaged activity” (7-item) referred to their perceived preference for and level of enrichment regarding the question-generation and peer-assessment tasks involved in science learning.

Items from existing validated instruments were adapted to fit the targeted experimental context. Specifically, when constructing the “Perception Toward Assessors Scale,” “Peer Assessment Rating Form” (Lurie et al., 2006), “The Questionnaire on Teacher Interaction” (Wubbels & Brekelmans, 2005), “Student Perceptions of their Own Dyad” (Yu, 2001), “Student Perceptions of Other Dyads” (Yu, 2001) and “Cooperative Learning Scale” (Ghaith, 2003) were referred to. Xiang’s (1979) “Classroom Environment Scale” and Li’s (1980) “Learning Environment Scale” were adopted and adapted for the “Student Perceptual Impressions of Classroom Climate.” Finally, “Learning Environment Dimensions in the CCWLEI Questionnaire” (Frailich, Kesner, & Hofstein, 2007) and “Clinical Learning Environment Scale” (Dunn & Hansford, 1997) were referred to for the “Perception toward the Engaged Activity Scale.”

Considering that the applied context and target audience of the referenced scales were different from those of the present study, all three scales were validated by another group of 193 5th graders (age 10-11) from nine classes from four different schools. Only items passing
item-total correlation, item-analysis and exploratory factor analysis (factors with eigenvalues greater than 1.0 and items with a targeted factor loading greater than 0.50) were included in the actual study. Details of each of the adopted scales are summarized in Table 1. As shown, the three scales adopted in this study evidenced good validities and reliabilities.
<table>
<thead>
<tr>
<th>Scale index</th>
<th>Dimensions with factor</th>
<th>Total variance explained</th>
<th>Sample items</th>
</tr>
</thead>
</table>
| **Perceptions toward their assessors** *(0.93)* | 1. The personality attributes of the assessors *(0.88~0.52)* | 67.69% | - My assessors cared about how much I learned;  
- My assessor was a responsible person;  
- My assessor could assist my learning;  
- My assessor cared about how I felt;  
- My assessor made me feel a sense of frustration;  
- My assessor did not possess enough knowledge with regards to the assessed item. |
|            | 2. Students’ interpretation of their relationship with the assessors *(0.85~0.71)* |               |              |
|            | 3. The competency of their peers as assessors *(0.80~0.74)* |               |              |
| **Classroom climate** *(0.81)* | 1. Perceived classroom order *(positive) *(0.80~0.69)* | 68.03% | - The majority of students could concentrate on the online question-generation and peer-assessment activity;  
- During the online question-generation and peer-assessment activity, the class was a bit noisy. |
|            | 2. Perceived classroom order *(negative) *(0.92~0.90)* |               |              |
| **Perceptions toward the engaged learning activity** *(0.93)* | Perceptions toward the engaged learning activity *(0.86~0.74)* | 67.55% | - The online question-generation and peer-assessment activity enhanced my interest in learning science;  
- I enjoyed participating in the question-generation and peer-assessment activity. |
3. Results

3.1 Background information on participants and descriptive statistics of examined variables

The demographic information of participants (N=243) are listed in Table 2. Approximately half of the subjects were males in each treatment condition. Less than 45% of the subjects had past experience in question-generation. The majority of the subjects (55%) did not have experience with peer-assessment. The potential effects of gender and prior experience on peer-assessment performance have been reported in previous studies (Topping, 2010). The $\chi^2$ results showed gender dispersion and subjects’ prior experiences in question-generation and peer-assessment to be invariant among all four treatment conditions in the study ($p = 0.97, 0.19, \text{ and } 0.42, \text{ respectively}$); thus, their potential effects on observed variables could be discounted.

Table 2: Background information for the four identity revelation modes

<table>
<thead>
<tr>
<th>Treatment groups</th>
<th>Demographic variables</th>
<th>Real-Name (n=64)</th>
<th>Anonymity (n=53)</th>
<th>Nickname (n=64)</th>
<th>Self-Choice (n=62)</th>
<th>$\chi^2$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>32</td>
<td>25</td>
<td>32</td>
<td>29</td>
<td></td>
<td>0.23</td>
<td>0.97(ns)</td>
</tr>
<tr>
<td>Female</td>
<td>32</td>
<td>28</td>
<td>32</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience in Question-Generation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.80</td>
<td>0.19(ns)</td>
</tr>
<tr>
<td>Yes</td>
<td>29</td>
<td>26</td>
<td>20</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>35</td>
<td>27</td>
<td>44</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience in Peer-Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.80</td>
<td>0.42(ns)</td>
</tr>
<tr>
<td>Yes</td>
<td>20</td>
<td>17</td>
<td>17</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>44</td>
<td>35</td>
<td>47</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: ns denotes non-significant

The means and standard deviations of pre- and post-assessment of the observed variables are listed in Table 3. Also included are the total number of questions generated and assessed during the study and some interaction indices, including the total number of feedbacks received from assessors, the total number of assessors with whom each student interacted and the number of interaction cycle each student engaged in. The interaction cycle was defined as the number of
feedbacks received from the assessors plus the number of responses the author of the question gave in respect to the received feedback per question. In total, 1753 questions were generated, and 63.15% of the questions were assessed (Table 3). ANOVA indicated no significant differences in the number of questions generated or assessed among the four groups ($p = 0.41, p = 0.12$).

**Table 3:** Descriptive statistics of observed variables of the four identity revelation modes

<table>
<thead>
<tr>
<th>Observed variables</th>
<th>Treatment groups</th>
<th>Real-Name (n=64)</th>
<th>Anonymity (n=53)</th>
<th>Nickname (n=64)</th>
<th>Self-Choice (n=62)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceptions toward assessors</td>
<td>Pre M (SD)</td>
<td>3.79(0.84)</td>
<td>3.23(0.72)</td>
<td>3.18(0.83)</td>
<td>3.45(0.74)</td>
</tr>
<tr>
<td></td>
<td>Post M (SD)</td>
<td>3.83(0.82)</td>
<td>3.14(0.89)</td>
<td>3.09(0.94)</td>
<td>3.57(0.81)</td>
</tr>
<tr>
<td></td>
<td>Adjusted M</td>
<td>3.64</td>
<td>3.25</td>
<td>3.21</td>
<td>3.55</td>
</tr>
<tr>
<td>Perceptions toward classroom climate</td>
<td>Pre M (SD)</td>
<td>3.47(0.8)</td>
<td>2.97(0.77)</td>
<td>3.07(0.81)</td>
<td>3.42(0.79)</td>
</tr>
<tr>
<td></td>
<td>Post M (SD)</td>
<td>3.44(0.81)</td>
<td>2.85(0.75)</td>
<td>2.96(0.94)</td>
<td>3.28(0.80)</td>
</tr>
<tr>
<td></td>
<td>Adjusted M</td>
<td>3.28</td>
<td>2.99</td>
<td>3.06</td>
<td>3.23</td>
</tr>
<tr>
<td>Perceptions toward the learning activity</td>
<td>Pre M (SD)</td>
<td>4.12(0.82)</td>
<td>3.58(0.93)</td>
<td>3.49(1.16)</td>
<td>3.71(0.98)</td>
</tr>
<tr>
<td></td>
<td>Post M (SD)</td>
<td>3.95(0.94)</td>
<td>3.57(0.98)</td>
<td>3.62(1.07)</td>
<td>3.78(1.01)</td>
</tr>
<tr>
<td></td>
<td>Adjusted M</td>
<td>3.71</td>
<td>3.69</td>
<td>3.76</td>
<td>3.76</td>
</tr>
<tr>
<td>Number of questions generated</td>
<td>Total</td>
<td>415</td>
<td>502</td>
<td>431</td>
<td>405</td>
</tr>
<tr>
<td>Number of questions assessed</td>
<td>Total</td>
<td>283</td>
<td>323</td>
<td>280</td>
<td>221</td>
</tr>
<tr>
<td>Number of feedbacks received from assessors</td>
<td>Total</td>
<td>841</td>
<td>831</td>
<td>761</td>
<td>748</td>
</tr>
<tr>
<td></td>
<td>Cognitive-substantial</td>
<td>89</td>
<td>168</td>
<td>167</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Cognitive-technical</td>
<td>662</td>
<td>554</td>
<td>602</td>
<td>441</td>
</tr>
<tr>
<td></td>
<td>Positive emotion</td>
<td>189</td>
<td>229</td>
<td>176</td>
<td>208</td>
</tr>
<tr>
<td></td>
<td>Negative emotion</td>
<td>99</td>
<td>108</td>
<td>115</td>
<td>127</td>
</tr>
<tr>
<td>Number of assessors with whom each student interacted weekly</td>
<td>Average$^b$</td>
<td>3.05</td>
<td>2.99</td>
<td>2.82</td>
<td>2.85</td>
</tr>
<tr>
<td></td>
<td>Pre M (SD)</td>
<td>3.04(3.53)</td>
<td>2.61(0.37)</td>
<td>2.87(3.93)</td>
<td>2.58(3.31)</td>
</tr>
<tr>
<td></td>
<td>Post M (SD)</td>
<td>2.78(2.41)</td>
<td>2.65(2.47)</td>
<td>2.73(2.97)</td>
<td>2.76(2.31)</td>
</tr>
<tr>
<td></td>
<td>Adjusted M</td>
<td>2.75</td>
<td>2.67</td>
<td>2.72</td>
<td>2.78</td>
</tr>
<tr>
<td>Number of interaction cycles each student engaged in per question</td>
<td>Average$^b$</td>
<td>4.73</td>
<td>4.02</td>
<td>3.49</td>
<td>4.28</td>
</tr>
<tr>
<td></td>
<td>Pre M (SD)</td>
<td>4.92(6.13)</td>
<td>3.09(3.27)</td>
<td>3.35(4.59)</td>
<td>3.67(6.37)</td>
</tr>
<tr>
<td></td>
<td>Post M (SD)</td>
<td>4.55(5.05)</td>
<td>4.07(4.06)</td>
<td>3.27(3.48)</td>
<td>3.89(3.68)</td>
</tr>
<tr>
<td></td>
<td>Adjusted M</td>
<td>4.34</td>
<td>4.19</td>
<td>3.34</td>
<td>3.90</td>
</tr>
</tbody>
</table>
Note. $M^* = \text{mean (on a five-point scale)}$

Average$^*$ refers to weekly interaction conditions.
Pre $M^*$ refers to the interaction condition of week 2, while Post $M^*$ refers to the week 6 interaction condition.
Adjusted $M^*$ refers to adjusted interaction condition during week 6, taking into account the covariate.

3.2 The effects of different identity revelation modes on perceptions toward assessors, classroom climate and the learning activity

As a result of the statistically significant moderate correlation coefficients among the three examined variables, ranging from 0.46 to 0.61, data was analyzed using multivariate analysis of covariance (MANCOVA) with student scores on the respective scale of the pre-session questionnaire used as covariates (Hair, et.al., 2006). The assumptions of normality and homoscedasticity were satisfied for each individual variable separately ($p=0.44$, $0.17$, $0.26$) and collectively ($p=0.13$). The result of the MANCOVA, using Roy’s largest root statistic, indicated that students in the real name, anonymity, nickname and self-choice modes performed statistically differently ($F= 6.93$, $p<0.01$, power =0.98).

Follow-up ANCOVA indicated that students in the different treatment-groups exhibited statistically different perceptions toward assessors ($F=4.88$, $p =0.003$, $\eta^2 = 0.06$). However, no evidence supported the premise that students in different identity revelation modes perceived statistically differently with regard to classroom climate ($F=2.11$, $p =0.10$, $\eta^2 = 0.03$), or the activity they were engaged in ($F=0.11$, $p =0.96$, $\eta^2 = 0.001$). A post-hoc comparison of perceptions toward assessors via Scheffé test, presented in Table 4, further indicated that participants in the real-name group viewed their assessors significantly more positively as compared to the anonymity and nickname groups ($p =0.006$, $0.002$, respectively). Additionally, students in the self-choice group perceived their assessors more positively as compared to the
anonymity \((p = 0.03)\) and nickname groups \((p = 0.01)\).

**Table 4**: Post-hoc comparisons among treatments on perceptions toward assessors

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Treatment (I)</th>
<th>Treatment (J)</th>
<th>Mean Differences (I-J)</th>
<th>SE</th>
<th>p-value</th>
<th>95% Confidence Interval Lower</th>
<th>95% Confidence Interval Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real-Name</td>
<td>Anonymity</td>
<td>0.39*</td>
<td>0.14</td>
<td>0.006</td>
<td>0.11</td>
<td>0.11</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Nickname</td>
<td>0.43*</td>
<td>0.13</td>
<td>0.002</td>
<td>0.17</td>
<td>-0.17</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Self-Choice</td>
<td>0.09</td>
<td>0.13</td>
<td>0.51</td>
<td>-0.17</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Anonymity</td>
<td>Nickname</td>
<td>0.41</td>
<td>0.14</td>
<td>0.77</td>
<td>-0.23</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-Choice</td>
<td>-0.30*</td>
<td>0.14</td>
<td>0.03</td>
<td>-0.57</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>Nickname</td>
<td>Self-Choice</td>
<td>-0.34*</td>
<td>0.13</td>
<td>0.01</td>
<td>-0.60</td>
<td>-0.08</td>
<td></td>
</tr>
</tbody>
</table>

3.3 **The pattern of identity mode use by nickname and self-choice groups**

To understand how students in the nickname and self-choice groups made use of the identity modes, an in-depth analysis of the identity use pattern of the two groups was conducted. Results evidenced that students in the nickname group used on an average of 3.2 nicknames after the designated treatment condition was implemented at the third week of the study. During the four week period, the frequency of nickname change ranged from 1 to 18, with the mode resting at 3. In-depth examination of the timing of nickname changes indicated that some students changed their nickname when assessing different questions even during the same instructional session, while others changed nicknames when they moved from the question-generation activity to the peer-assessment activity on a weekly basis.

Additional analysis done on the identity revelation method chosen by the self-choice group revealed that more than 90% (96% to be exact) of the participants made use of at least two available identity modes at different occasions during the study, of which 56% tried all three available identity modes during the study.
3.4 Interaction analysis

As shown in Table 3, the average number of assessors with which each question-author interacted on a weekly basis ranged from 2.82 to 3.05. The average interaction cycles between the question-author and the assessor ranged from 3.49 to 4.73 per question. Further ANCOVA examining differences among the four identity revelation modes evidenced a similar intensity of interaction among the four modes (p=0.99, p=0.60).

To have a better understanding of the nature of the feedback provided by assessors, a content analysis of comments among interacting parties was conducted using an interaction process analysis method, followed up by ANCOVA. The framework proposed by Guan, Tsai and Hwang (2006) was referred to for the interaction process analysis, and the categories that bore the most relevancy to the study were adopted and adapted. Explicitly, comments offered by assessors were analyzed along four dimensions: cognitive dimensions on substantive areas (i.e., science content), cognitive dimensions on technical areas (i.e., question-generation), positive emotional content and negative emotional content.

Several results were obtained. First, no statistically significant differences in the frequency of occurrence were found in either the cognitive dimensions (substantive or technical areas) or the negative emotional dimension among the four treatment groups (p=0.054, p=0.19, p=0.24, respectively). Despite this, the severity level of the negative comments given by the assessors from the anonymous and nickname groups, as compared to the real-name group, was evident. Particularly, comments that implicated intense, impulsive and irrational emotions were present in both anonymous and nickname groups, as expressed in words like “POOR, POOR, POOR!”;
idiot!; ridiculous!; good for nothing!!”. Second, a statistically significant difference was found ($p=0.002$) in the positive emotional feedback provided by assessors. Post-hoc comparisons via Scheffé test further showed that assessors in the self-choice group offered significantly more positive emotional comments than those in the nickname group ($p=0.001$).

4. Discussion and Conclusions

Due to media differences and affordances innate with anonymity, real-name and nickname situations and prior research witnessing students exhibiting significant different preferences toward identity revelation modes (Yu & Liu, 2009), the researchers in this study followed up by examining if there were any differential effects among the four different identity revelation modes (three fixed modes: real name, anonymity, nickname and one dynamic user self-choice mode) on participants’ perceptions toward their assessors, their classroom climate, and the learning activity in which they were engaged as evidenced by a self-reported questionnaire. The current study confirmed that different identity modes led participants to view their assessors differently. Specifically, participants assigned to the self-choice identity revelation mode tended to view their assessors more favorably than those in the anonymity and nickname groups. Furthermore, participants in the real-name group also viewed their assessors more favorably than those in the anonymity and nickname groups. However, different levels of identity revelation did not lead participants to view their learning climate or the learning activity differently.

Several additional results regarding the interaction process between assessors and the question-author, contributing an explanation of the above-mentioned results, are discussed herein. First, as suggested by the deindividuation theory, a severe level of negative comments,
implicating intense, impulsive and irrational emotions were present in both the anonymous and nickname groups, which might be a result of the assessor’s unrecognizable or less recognizable identity and a decreased sense of individual accountability. The fact that any such severe negative comments, no matter how rarely they happened, once received by students in the anonymous and nickname groups, accounted for less overall favorable ratings towards assessors, as observed in the study, is understandable. Second, in the case in the self-choice group, the combined effects of granting students the freedom to choose whichever identity revelation mode they preferred and the overall more positive emotional feedback they received from assessors may, as a result, have created a more favorable attitude toward their assessors.

Third, the identity use pattern regarding the frequency and timing of nickname changes as presented in the previous section was indicative of students in the nickname group intending to disguise their identity from their peers, which resulted in the nickname condition being close to the anonymity condition. This similar degree of “un-identifiability” of both the nickname and anonymity groups helped to explain why these two groups in the study arrived at similar results in most of the observed variables.

Finally, the authors conjectured that different identity revelation modes might result in participants viewing the learning climate and the learning activity differently; however, data analysis fails to reject the null hypotheses. In other words, the identifiability of all participants in the real-name group and some in the self-choice or nickname groups neither led to a presumably chaotic, ill-managed classroom, nor did it lead to less favorable attitudes toward the activity that was introduced. In fact, students in all treatment groups appeared to be quite focused on the task assigned to them. This informal observation was supported by the similar number of interaction
cycles and the similarity in the intensity of the interaction between the question-author and assessors as discussed in the result section. Therefore, the equivalent level of engagement in the learning tasks (question-generation and peer-assessment), may very well have led to all the four treatment groups perceiving the classroom climate and the activity at similar levels and may, in turn, have left no effects from the different identity revelation modes.

4.1 Significance and Implications of the study

The obtained findings have important empirical significance as well as implications for online system development, instructional implementation and future study. First, this is the first study examining the effect of a self-choice identity revelation mode with regard to interpersonal relationships in peer-assessment situations and the first to show significantly more favorable perceptual impressions of question-authors toward their assessors in a self-choice mode, as compared to the anonymity and nickname groups.

Second, while previous studies have found that a predominante percentage of participants opt for anonymity (37.5%) or nicknames (35%), with considerably fewer respondents leaning toward a real-name mode (only 12.5%) as their preferred identity mode when providing comments to their classmates (Yu & Liu, 2009), anonymity and nickname modes did not lead to significantly more positive ratings toward assessors, learning climate or the learning activity. What’s more surprising is that participants in the real-name group (similar to those in the self-choice group) viewed their assessors significantly more favorably than those in the anonymity and nickname groups. In other words, as evidenced by the present study, arbitrarily imposing either nicknames or anonymity modes negatively affects learner perceptions toward their assessors as compared to
real-name or self-choice modes. This study directs the attention of both researchers and instructors to the potential adverse effects of arbitrarily imposing anonymity or nickname modes in peer-assessment situations.

Third, while more and more online learning systems embed the elements of anonymity and nicknames in order to supplement the real-name identity revelation mode, almost all of them currently adopt a fixed mode rather than a dynamic mode that is adjustable in time and by the user. Based on the present study in which self-choice and real-name identity modes were found to result in participants viewing their assessors more favorably than was the case with anonymity or nickname modes, designers of online systems should consider building a self-choice identity revelation mode into their online systems so as to accommodate human intrinsic motivation for control (Malone & Lepper, 1987). As for instructors interested in online peer-assessment, they are advised to opt for systems built with self-choice identity or real-name modes if interpersonal relationships in particular are what they are aiming for at the time.

Fourth, in this study, participants’ perceptions toward their assessors were found to be significantly different among the four different treatment groups. However, further data analysis on system use and a content analysis of online interaction history indicated no significant difference in any of the following indexes: the total number of questions generated and assessed, the total number of cognitive and negative emotional comments in the assessor’s feedback, and so on. Despite the fact that additional data analyses on system use combined with those done on the questionnaire do not yield to a consistent pattern in this study, this points to an interesting phenomena: actual system use does not necessarily literally amount to personal subjective opinion. In light of the findings of Wilson (2004), students’ attitudes toward certain aspects of a
system (i.e., attitudes concerning usefulness and importance) are more important than actual system use in predicting grades, which highlights the importance of participants’ attitudes rather than the more frequently regarded objective system use. Therefore, instructors and researchers alike are advised not to merely depend on readily available actual system use while ignoring or de-emphasizing users’ attitudes and perceptions with regards to their own experiences, as afterwards, it is student subjective construction of the reality to which they are exposed that may be more relevant.

### 4.2 Generalizations of the study

Researchers and practitioners are advised to note that the current study was conducted in primary school settings, and as such, generalizations of the obtained findings to other contexts (for instance, different age groups, different learning activities, and so on) should be exercised with care.

**Acknowledgements**

This paper was funded by a 3-year research grant from the National Science Council, Taiwan, ROC (NSC 96-2520-S-006-002-MY3).

**References**

Studies 61(1), 84-103.


Psychology, 33, 178-183.


Ghaith, G. (2003). The relationship between forms of instruction, achievement and perceptions of


