Online Targeting Behavior of Peer-Assessors under Identity-Revealed, Nicknamed, and Concealed Modes

Fu Yun Yu and Shannon Sung

1Institute of Education, National Cheng Kung University, Taiwan // 2Education Studies Program, Spelman College, USA // fuyun.ncku@gmail.com // ssung@spelman.edu

Corresponding author

(Submitted August 4, 2016; Revised March 8, 2017; Accepted August 2, 2017)

ABSTRACT

This study examined whether different identity revelation conditions result in different online targeting behavior among peer-assessors through a pretest and posttest quasi-experimental research design. Students from six fifth-grade classes (N = 196) participated in online learning tasks where they generated and selected peer-generated questions to review and assess in their respective identity revelation modes—real-name, nickname, and anonymity. Several findings were obtained. First, there was a high redundancy rate for both high- and low-targeted assessees under the pre-treatment (control) and treatment conditions for all three identity revelation modes. Second, the non-significant results of the chi-square tests indicated that the identity revelation modes and assessees redundancy were neither considerably related for the high- nor low-targeted assessees groups. Third, the results from the ANCOVA on the number of peer-feedback messages each student received showed no significant differences among the three identity revelation treatments. Fourth, the Wilcoxon tests confirmed that there were no significant differences in the assessees rankings in terms of the number of times their generated questions were assessed between the pre-treatment and treatment conditions for all different identity revelation modes. Finally, the results from participants’ responses to the checkbox question revealed almost the same ranking pattern regarding factors dominating their targeting behavior (with the question-author ranked last among all factors), despite the various different identity revelation modes. In sum, identity revelation modes were not found to affect peer-assessors’ targeting behavior in an online peer-assessment activity.

Keywords

Identity, Online targeting behavior, Peer-assessment

Introduction

Assessment is one of the decisive factors affecting the processes and outcomes of teaching and learning. Conventional teacher-centered methods of assessment are now giving way to alternative and innovative assessments that enable the learnes to actively share responsibility and initiative in assessing personal levels of understanding and growth (Cheng & Warren, 2005). Understandably, the potential of peer-assessment has attracted much attention since its emergence and has been found to promote critical thinking, cognitive restructuring, better quality work, learning motivation, self-efficacy, and attitudes toward learning (Cartney, 2010; Cheng & Warren, 2005; Kaufman & Schunn, 2011; Topping, 2009; van Gennip, Segers, & Tillema, 2010).

Recent developments in peer-assessment research

Recently, the effects of different types and arrangements of peer-assessment have been the focus of many studies, for example, anonymous vs. identified peer-assessment (Yu & Wu, 2011), concise vs. elaborated peer-feedback (Srihobos, Narciss, & Dünnheimier, 2010), peer-feedback with vs. without justification (Bolzer, Strijbos, & Fischer, 2015), non-directive vs. specific peer-feedback (Cho & MacArthur, 2010), corrective peer-feedback with different degrees of directness (AbuSeileek & Abualsha’r, 2014), and peer-feedback with different levels of structure (Gielen & De Wever, 2015). Studies concerning the effects of the degree of anonymity of the assessor and assessed have special importance since researchers (e.g., Ballantyne, Hughes, & Mylonas, 2002; White, 2009) have questioned the objectiveness and fairness of peer-assessment. For instance, the outcomes of peer-assessment have been reported to be based more on the personal views of the assessed rather than the work being examined, with more extrovert individuals usually receiving higher scores than those who are more introverted (Johnston & Miles, 2004). Friendship has also been found to influence assessments, where individuals belonging to the same inner circle receive inflated scores as opposed to those outside such circles (Chan, 2010) or when there is a high level of friendship between the assessors and assessees (Panadero, Romero, & Strijbos, 2013).

Because the disclosure and concealment of identities between interacting parties in a peer-assessment situation could be easily managed in online spaces (Yu, 2009), researchers have examined its effects on various
educational outcomes, such as the quality of the work produced, interpersonal relationships, perceptions toward peer-assessment, ratings and comments provided, professional behavior, and engagement (Ainsworth, Gelmini-Hornsby, Threapleton, Crook, O’Malley, & Buda, 2011; Garner, McKendree, O’Sullivan, & Taylor, 2010; Raes, Vanderhoven, & Schellens, 2015; Li, 2017; Stepanyan, Mather, Jones, & Lusuardi, 2009; Thompson & McGregor, 2009; Wadhwa, Schulz & Mann, 2006). Motivated by the fact that few studies have investigated its effects on the actual interaction “process” or related behavior, the online targeting behavior of the assessor (i.e., decision-making as to which item to assess) was examined in one prior study using a one-group research design to understand the “process” of interaction in identified (pre-treatment) and anonymous (treatment) peer-assessment situations (Yu & Sung, 2016).

Given that the aforesaid design could have been weakened by several internal validity threats, such as history and maturation (Grinnell & Unrau, 2010), a more robust research design—a pretest-posttest quasi-experimental research method, was adopted in this study. Furthermore, since researchers suggest that different levels of user identifiability may lead to different interactions (Yu & Wu, 2011), and the use of nicknames is prevalent among online users, the online targeting behavior of the assessor using real-name (i.e., absolute revelation), nickname (i.e., partial revelation/concealment), and anonymity (i.e., complete concealment) modes is examined in this study.

Research questions

The assessors’ actual behavior when selecting a target for online peer-assessment (i.e., online targeting behavior) and their reasons for such behavior in three identity revelation modes were investigated. The two main research questions formulated for the study are as follows:

RQ1: Do different identity revelation modes (i.e., real-name, nickname, and anonymity) among the interacting parties affect the online interaction behavior and social dynamics of the group in online peer-assessment contexts? To address this, four sub-research questions are examined:

- Does the proportion of overlapping assessees among the high-targeted assessees (i.e., assessees who receive a comparatively high number of peer-feedback messages) vary significantly among the three different identity revelation modes?
- Does the proportion of overlapping assessees among the low-targeted assessees (i.e., assessees who receive a comparatively low number of peer-feedback messages) vary significantly among the three different identity revelation modes?
- Does the average number of peer-feedback messages each student receives differ among the three identity revelation treatments?
- Does ranking in terms of the total number of peer-feedback messages received in different identity revelation modes differ significantly between the pre-treatment (i.e., control) and treatment conditions?

RQ2: Do the major reasons selected by peer-assessors for their online targeting behavior vary significantly among the three different identity revelation modes?

Literature review

Face-to-face peer-assessment

Peer-assessment capitalizes on the idea of peers assisting one another by providing and receiving quantitative and/or qualitative feedback to and from their equivalent peers about their respective work or performance (van Gennip et al., 2010). Because peers generally have a similar learning status and are within each other’s zone of proximal development, peer-feedback is more easily perceived and understood, as compared to that given by teachers (Topping, 2009). Studies have found that students’ self-esteem and self-efficacy regarding academic competence tend to be more positively affected by evaluations from peers than from their teachers (Kaufman & Schunn, 2011), and the quality, validity, and reliability of peer-assessment has been attested in a number of works (e.g., Avery, 2014; Jones & Alcock, 2014).

Although the benefits of peer-assessment have been recognized by teachers, its adoption in classrooms has not been totally encouraging. Foremost, face-to-face peer-assessment is spatial- and physically-bound, as well as time-consuming and labor-intensive regarding material preparation and data compilation (DiPardo & Freedman, 1988). Peer-assessment may also be in conflict with the societal norms in certain cultures that implicitly
discourage confrontational discourse or critical feedback (Cartney, 2010), and this may result in inaccurate assessments or peer pressure especially in identified face-to-face contexts (Raes et al., 2015). Finally, students may not be equipped with the required capacity to provide constructive feedback to their peers without adequate scaffolds to ensure quality peer-assessment work (Liang, 2010). For all these reasons, many researchers have turned to computer-mediated peer-assessment.

**Computer-mediated peer-assessment**

The affordances of networked technologies are crucial in alleviating the challenges of face-to-face peer-assessment, while at the same time enabling authors to receive timely and constructive feedback. Many online peer-assessment systems have been developed in the past decade to help manage and implement peer-assessment effectively in classrooms (e.g., NetPeas, Vee heuristic, Web-SPA, and SWoRD). With unique features including high processing speed, immense storage space, learner control of sequencing, customizable scaffolds for feedback-provision, multimedia, simultaneity, instantaneity, and space-, time- and device-independence, various technological and pedagogical arrangements can be made available to online users of such systems (Yu & Wu, 2011). System features, such as automatic assignment of work to be assessed, a set of customized built-in peer-assessment criteria for reference, time-stamping, multiple peer reviews, asynchronous and synchronous interaction, anonymity, process display, history record, and real-time notification are frequently embedded in these systems to expedite and streamline tasks (Smaldino, Lowther, & Russell, 2008; Yu & Wu, 2011).

Although online peer-assessment generally benefits students cognitively and affectively, it has several shortcomings, including difficulties related to comprehension because it lacks contextual cues, which may compromise its effects (Jones & Alcock, 2011; Liang, 2010). In an attempt to rectify this and further promote the learning effects of peer assessment, many researchers have focused on exploring different peer-assessment designs (e.g., anonymous/identified peer-assessment, concise/elaborated peer-feedback, peer-feedback with/without justification, non-directive/specific peer-feedback, and corrective peer-feedback with different degrees of directness), and examine their relative effects.

The psychological functioning of anonymity during online interactions—social identity theory and sense of perceived psychological safety

The social identity theory describes how people develop and form a sense of belongingness, membership, and social identity by associating themselves with particular affiliations (Pearce II, 2013; Tinson & Close, 2012). In addition to forming personal identities as individuals, people tend to search for and establish their social identities through group affiliation, which differentiates them from other groups (Hogg, 1987). This theory has been used to explain and predict how perceived status differences and personal characteristics in social groups govern an individual's behavior and decision-making (Hogg, van Knippenberg, & Rast III, 2012; Tajfel & Turner, 1986). Accordingly, individuals perceived as having a prominent social identity (i.e., being superior or more favored) are believed to possess more reliable sources of information and thus will tend to exert more influence over the perceptions and behavior of other group members (Hogg et al., 2012). In contrast, those perceived to have a less prominent social identity are less likely to receive as much attention as those from more favored and prominent groups (Tinson & Close, 2012). Considering the potential for prejudiced interactions, it has been suggested that concealing the real identities of participants during group interactions could avoid judgments being formed on the sheer basis of the person’s affiliations and could help foster greater and more balanced participation among individuals of different backgrounds (Chester & Gwynne, 1998).

On the other hand, anonymity’s potential to remediate biased scoring and differential participation is believed to be associated with an enhanced sense of psychological safety (Yu, 2009; Miyazoe & Anderson, 2011; Roberts & Rajah-Kanagasabai, 2013) — a state where one feels secure and comfortable without excessive concerns over any possible effects on one’s well-being, image, and status (Zhang, Fang, Wei, & Chen, 2010). Identity concealment could help lessen negative emotions, such as apprehension, intensified pressure, and a sense of insecurity (Bullingham & Vasconcelos, 2013), which may be caused by social comparison, peer pressure, and self-validation (Franzoi, 2006; Moral-Toranzo, Canto-Ortiz, & Gómez-Jacinto, 2007; Raes et al., 2015). Similarly, excessive self-consciousness, which is linked closely with identifiability in group situations and may inadvertently affect assessors’ objective feedback, can be reduced by masking one’s identity (Bullingham & Vasconcelos, 2013; Wadhwa et al., 2006). Overall, psychologically safe learning contexts are generative and have been found to be associated with more active knowledge-sharing and genuine self-expression intention (Zhang et al., 2010), which can lead to improved work or more refined performance (van Gennip et al., 2010).
The psychological functioning of nicknames in online interaction situations—Identity construction

Nicknames disguise the real identity of participants, and may thus have similar liberating effects on the interacting parties as anonymous situations do. As such, the social identity theory and psychological safety serve well as the theoretical foundations for the inclusion of a nickname option in virtual worlds. Nonetheless, the use of nicknames may engender further benefits, as reported in identity construction studies.

Studies on identity construction stress the importance of one’s identity formation through reciprocal interaction with sociocultural factors (Bullingham & Vasconcelos, 2013; Hassa, 2012). Such studies mostly highlight the social constructive process of self-perception that is closely intertwined with language, culture, and society (Krämer & Haferkamp, 2011; Locher, 2008) and transcends conventional social categorizations (e.g., gender, Cornetto & Nowak, 2006), which is more easily attainable in online spaces.

In the cyberworld, interpersonal communications and conversations are extended (i.e., not limited to face-to-face contexts) (Tufekci, 2008), and it is possible for members to use nicknames, profile photos, personas, avatars, and usernames, where one’s identity can be explored and formed (Lindholm, 2013; Subrahmanyam & Šmahel, 2011). Indeed, people form impressions of others based on users’ online nicknames (Chester & Gwynne, 1998), which can trigger interpersonal interactions (Lindholm, 2013). As emphasized by personality development theorists (e.g., Marcia, 1966), it is essential to allow individuals to be able to actively explore options before reaching their final stage of identity development. Enabling users to recreate certain entities or characteristics enables them to avoid value-laden contradictions or confrontations or can even help them impress others and attract attention (Krämer & Haferkamp, 2011). In sum, the representation of oneself in the virtual world provides opportunities for identity exploration and construction in a risk-free fashion (Subrahmanyam & Šmahel, 2011).

To summarize, peer-assessment learning activities are frequently in the center of online teaching and learning designs. Most studies focus on the effects of different types and arrangements of peer-assessment; however, some issues remain unresolved. In particular, among the studies examining identity revelation, very few focus on the participants’ actual interaction behavior, and far fewer involve the use of nicknames. Since selecting the targets to be assessed is the first step in the peer-assessment process and may impact student experiences, attitudes, and/or learning outcomes, and because the allocation of assessed work is commonly practiced in peer-assessment situations, issues surrounding whether or not revealing the identity of the author may reduce the chances that the participants’ work will be assessed still need to be resolved before instructors can confidently apply self-selecting targets for peer-assessment. Thus, the effects of real-name, nickname, and anonymity on peer-assessors’ actual online targeting behavior and their accompanying reasons for targeting any specific items were investigated in this study.

Methods

Participants and the online learning system

Six fifth-grade classes (age 10-11, N = 196, male = 96, female = 100) were randomly selected from a primary school in Taiwan and introduced to an online question-generation and peer-assessment learning activity during their weekly computer literacy class. Computer literacy is a required course from third grade on, so the participants all had fundamental computer capabilities (e.g., keyboarding, web-searching, using a word-processor, inserting multi-media files into a project, and so on). An online learning environment—Question-Authoring and Reasoning Knowledge System (QuARKS) was adopted to support the focal activity, which was only accessible in class to ensure that all the participants had the same amount of time to use the system. Even though QuARKS is capable of customization (in terms of the specific sets of functions accessible to users in different contexts) (Yu, 2009), for the purpose of this study, with the exception of different identity revelation modes that were assigned to different treatment groups, the accessible functions and implementation procedures were kept identical for all treatment groups.

Specifically, the only difference in the treatment groups was whether the real-name or nickname of the question-author was revealed or concealed in the question-author field. Explicitly, when the questions to be assessed were listed in the peer-assessment list window (see Figure 1) and could be viewed in the top portion (i.e., targeted question) of the online peer-assessment form window (see Figure 2), the student’s full-name was retrieved automatically from the database and shown in the question-author field in the respective windows for the real-name mode (Figures 1a and 2a). In contrast, for the nickname mode, the nickname used by the question-author
was shown (Figures 1b and 2b), and for the anonymity mode, no information about the identity of the question-author was given, and the term “anonymity” was shown instead (Figures 1c and 2c).

Figure 1. Peer-assessment list window for the real-name (left, a), nickname (middle, b) and anonymity (right, c) modes

Figure 2. The top portion of the online peer-assessment form window for the real-name (left, a), nickname (middle, b) and anonymity (right, c) modes

Figure 3. The bottom portion of the online peer-assessment form window (for the real-name mode)

While the default item-sequencing is based on the question submission time (from the latest, marked as question No. 1, to the earliest), to assist viewing and selecting targets, participants can change the order of items that appear on the system, according to the number of assessors/peer-feedback received, question type, question, and question-author, by clicking on the menu bar of the peer-assessment list window (Figure 1). Once students
decide their assessment target and click on the item of their choice (Figure 1), they are directed to the online peer-assessment form window to view the targeted question (Figure 2). They then provide both quantitative and qualitative feedback in the online peer-assessment form (Figure 3).

**Experimental design and implementation procedures**

To examine the effects of different identity revelation modes (i.e., real-name, nickname, and anonymity) on participants’ targeting behavior in an online peer-assessment context, a pretest-posttest quasi-experimental research design was adopted. The six participating classes were randomly assigned to the three different treatment groups (two classes in each mode).

Before the study, the students were briefed about the purposes of the online learning activity — supporting science learning and promoting higher order thinking skills, and informed that their performance on the tasks would account for part of their computer literacy grades. They then received a training session in the computer lab, where the following topics were covered: basic principles for question item writing, the question-generation criteria and models, the peer-feedback criteria and models, and operational procedures for question-generation and peer-assessment on QuARKS.

Afterwards, in the actual study during the 40-minute instructional time, the students were instructed to first construct two multiple-choice questions on QuARKS related to the science content covered in the current week. They were then directed to assess four peer-generated questions of their choice on QuARKS, where their quantitative ratings of the quality of the assessed item as well as descriptive feedback would be given and directed to the question-author for consideration. The students were asked to focus on their individual tasks and to not talk to their classmates. Any questions were thus directed to the implementer.

To establish a baseline of the students’ online targeting behavior, the real-identity mode was set for all experimental groups as the pre-treatment (control) condition for the first session, and the respective treatment conditions were introduced for the following week. In other words, the learning tasks completed by the participants in the three treatment groups for the pre-treatment and treatment conditions were the same — generating two questions and assessing four peer-generated questions. The only difference was in the identity revelation mode used for the question-author during the treatment condition. Furthermore, to enhance question-generation and peer-feedback performance, the students’ work was reviewed by the implementer, and three pieces from the respective classes were selected and shown as exemplars at the beginning of the treatment condition week for whole-class feedback.

Finally, the students responded to one checkbox question after the activity, where possible reasons that might have affected their targeting behavior during the online peer-assessment phase of the activity were provided for selection. A flow diagram summarizing the experimental procedures of this study is given in Figure 4.
Instrumentation

One checkbox question was developed and disseminated for individual completion after the study to collect data on the participants’ accounts of their online targeting behavior. Based on informal, non-participant observation of the participants’ selection behavior during the activities, six responses were listed as options to choose from — question type, number of assessors already giving feedback on the item, question item itself, question submission time, who the question-author was, and the current average rating. Participants were instructed to select not more than three reasons that they felt often directed their online targeting behavior during the peer-assessment activities in this study.

Data preparation and analysis

Prior to a statistical analysis, careful preparation of the data by applying appropriate filtering techniques is essential for accurate data analysis (Hansen, Shneiderman, & Smith, 2010). Specifically, to ensure an unbiased comparison between the pre-treatment (i.e., real-name mode, 1st week) and respective treatment conditions (i.e., different identity revelation modes, 2nd week), only individuals who participated in both activities (i.e., represented by \( n' \)) were retained for data analysis for research question 1. As a result, the number of assessees decreased considerably (where \( n' \) ranges from 14 to 23, \( \bar{n'}=19.2 \)) after deletion, as compared to the number of students in the original class, \( n \), ranging from 31 to 34, \( \bar{n}=32.7 \) (see Table 1).

To answer research questions 1-1 and 1-2, firstly, high- and low-targeted assessees were operationally defined. With reference to the literature on item discrimination indices (e.g., Beuchert & Mendoza, 1979; Engelhart, 1965) and taking into consideration the small sample size of this study, 33% was chosen as the upper and lower cutoff percentage. The high-targeted subgroups consisted of participants whose total number of received peer-feedback messages in the pre-treatment control condition and the subsequent treatment condition, respectively, were in the upper 33% of their respective treatment groups, whereas the low-targeted subgroups were those whose total number of received peer-feedback messages were in the lower 33%. The recurring members found in both the pre-treatment and treatment conditions were considered redundant assessees. The proportions of redundant assessees for the three different treatment groups were then calculated, and the results were analyzed using a chi-square test. The assessors’ online targeting behavior in different identity modes could be revealed in this way by factoring in the patterns of redundant assessees in both the high- and low-targeted subgroups.

To answer research question 1-3, the two classes receiving the same treatment were combined for statistical tests (Table 1). An ANCOVA was performed to determine if there were significant differences in the average number of peer-feedback messages each student received among the three identity revelation groups. To control for pre-existing differences, data on the same variable in the pre-treatment were used as covariates.

To answer research question 1-4, the two classes receiving the same treatment were combined and then ranked based on the total number of peer-feedback messages received in the pre-treatment condition and the subsequent treatment condition, respectively. The Wilcoxon test was adopted to reveal any significant ranking differences as a result of changes in the three identity revelation modes.

Finally, to answer research question 2, the reasons that participants assigned to the same treatment group chose for their online targeting behavior were tallied together and compared among different identity revelation modes.

Results

Research question 1: Do different identity revelation modes among the interacting parties affect the online interaction behavior and social dynamics of the group in online peer-assessment contexts?

To answer research question 1-1, a two-way contingency table analysis was conducted. The proportions of redundant assessees for the high-targeted subgroup were calculated first, and 68.8%, 66.7%, and 72.7% redundancy rates were found for the anonymity, nickname, and real name conditions, respectively (see Table 1). The results of the chi-square test were not significant, \( \chi^2 (2) = .110, p = .946 \), Cramér’s \( V \) high = .051, indicating that the identity revelation mode and asseseer redundancy were not significantly related for the high-targeted assessees subgroup.
Similarly, a two-way contingency table analysis was conducted to answer research question 1-2. The proportions of redundant assessees in the low-targeted subgroup were calculated first, for which a 66.7%, 63.6%, and 65.9% redundancy rate was found for the anonymity, nickname, and real name conditions, respectively (see Table 1). The results of the chi-square test were also not significant, $\chi^2 (2) = .033$, $p = .984$, Cramér’s $V_{low} = .028$, meaning that the identity revelation mode and assessee redundancy were not significantly related for the low-targeted assessees subgroup, either.

Furthermore, the assumption of the homogeneity of regression slopes was satisfied, $F (2, 109) = 1.267$, $p = .286$, before proceeding to the ANCOVA. The results of the ANCOVA conducted on the number of peer-feedback messages each student received among the different identity revelation treatments revealed no significant differences, $F (2, 111) = .186$, $p = .830$ (partial $\eta^2 = .003$).

Finally, for research question 1-4, the Wilcoxon test was performed, and the results showed non-significant differences in matched-pair signed rank between the pre-treatment and treatment conditions for all three different treatment groups (see Table 1).

<table>
<thead>
<tr>
<th>Identity revelation modes</th>
<th>Class</th>
<th>$n$ ($n^1$)</th>
<th>$m_{pre}$ (sd$_{pre}$)</th>
<th>$m_{post}$ (sd$_{post}$)</th>
<th>$m_{adj}$ (se)</th>
<th>Wilcoxon test</th>
<th>Redundant</th>
<th>Pearson $\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anonymity</td>
<td>1</td>
<td>34 (23)</td>
<td>2.04 (1.41)</td>
<td>2.20 (1.99)</td>
<td>2.32 (0.38)</td>
<td>Z = -.323</td>
<td>.68 66.7 0.11 0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>31 (22)</td>
<td>2.30 (2.07)</td>
<td>2.57 (2.71)</td>
<td>2.62 (0.40)</td>
<td>Z = -.707</td>
<td>.66 63.6 0.47 0.74</td>
<td></td>
</tr>
<tr>
<td>Nickname</td>
<td>3</td>
<td>32 (16)</td>
<td>2.26 (2.18)</td>
<td>2.83 (2.92)</td>
<td>2.59 (0.42)</td>
<td>Z = -1.482</td>
<td>.72 65.9 0.47 0.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>34 (22)</td>
<td>3.26 (2.18)</td>
<td>3.11 (2.82)</td>
<td>2.59 (0.42)</td>
<td>Z = -1.482</td>
<td>.72 65.9 0.47 0.74</td>
<td></td>
</tr>
<tr>
<td>Real-name</td>
<td>5</td>
<td>33 (18)</td>
<td>3.26 (2.18)</td>
<td>2.83 (2.92)</td>
<td>2.59 (0.42)</td>
<td>Z = -1.482</td>
<td>.72 65.9 0.47 0.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>32 (14)</td>
<td>3.26 (2.18)</td>
<td>2.83 (2.92)</td>
<td>2.59 (0.42)</td>
<td>Z = -1.482</td>
<td>.72 65.9 0.47 0.74</td>
<td></td>
</tr>
</tbody>
</table>

Note. $^1$ = the redundant participants; $^2$ = combined mean and standard deviation for pre-treatment condition; $^3$ = combined mean and standard deviation for treatment condition; $^4$ = adjusted mean and standard error.

Research question #2: Do the major reasons selected by peer-assessors for their online targeting behavior vary significantly among the three different identity revelation modes?

As seen in Table 2 and Figure 5, the rankings for the reasons for peer-assessors’ online targeting behavior in the three identity revelation modes reflected very similar patterns. First, the rankings were identical for the real-name and anonymity groups. Second, “question type,” “number of assessors already given feedback on the item,” and “the current average rating” were the three main reasons for all three identity revelation modes, with an average of more than 50% of all participants highlighting their influence on targeting. Third, “who the question-author is” was found to be the factor least likely to affect targeting behavior among the six listed reasons, preceded by “question item itself” and “question submission time” in that order.

Figure 5. Rankings of reasons for peer-assessors’ online targeting behavior among different identity revelation modes
Discussion and conclusions

This work investigated the targeting behavior of peer-assessors in three different identity revelation modes (real-name, nickname, and anonymity) in an online peer-assessment context. The probable reasons for peer-assessors’ online targeting behavior were also examined. Both social identity theory and psychological safety studies provided the basis for our proposed hypotheses regarding differences in targeting behavior in the identified and unidentified (i.e., anonymity and nickname) modes. Identity construction studies further provided support for the use of nicknames in online spaces.

Contrary to the authors’ hypotheses, the results from the inferential analyses all found no evidence supporting our proposed hypotheses, and all results pointed to the conclusion that changes in identity revelation modes did not significantly affect peer-assessors’ decisions as to which questions to assess in online peer-assessment situations. To summarize the major findings, foremost, both high- and low-targeted assesssee redundancy rates were high (exceeding 60%) for all of the experimental groups. The non-significant chi-square tests results further demonstrated that the identity revelation mode and assesseee redundancy were not significantly related in the high- or low-targeted assesssees groups. Moreover, the ANCOVA results on the number of peer-feedback messages each student received among the three conditions were non-significant. Finally, the non-significant differences in the Wilcoxon matched-pair signed rank tests reinforced the finding suggesting that the identity revelation mode of each treatment group did not significantly affect their participant rankings in terms of the total number of feedback messages received from peers.

The findings from the quantitative results could be explained and somewhat understood by the data collected from the end-of-session checkbox question. As revealed in Table 2 and Figure 5, the assessors made their decision on which questions to assess based predominately on “question type,” “number of assessors already given feedback on the item,” and “the current average rating.” They paid little attention toward who authored the questions when it came to selecting their targets to complete the online peer-assessment task. The non-significant effect results from the chi-square, ANCOVA, and Wilcoxon tests of all treatment groups can thus be understood.

To conclude, this study did not find supportive, empirical data to validate that masking real identities or recreating identities in an online context would induce different interaction patterns and behavior, specifically, targeting behavior, on the part of participants. In other words, the suspected impacts gleaned from interpersonal relationships and psychological functioning did not undermine the peer-assessors’ targeting behavior in this study. This study provided some initial empirical evidence to support past findings of very limited existing studies that suggest that being in an identifiable or anonymous situations makes no difference on the actual interaction process (Roberts & Rajah-Kanagasabai, 2013; Yu & Sung, 2016; Yu & Wu, 2011).

Implications of the study

This work has some implications for teaching practice and computer-supported learning system development. The non-significant results of the redundancy rate for high- and low-targeted assesssees in this study provide some comfort to elementary school instructors interested in introducing peer-assessment tasks in their classes, yet having no access to systems supporting anonymity or nickname use. As found in this study, the revelation or concealment of identity is unlikely to have any bearing on the number of times a focal item is going to be assessed.

<table>
<thead>
<tr>
<th>Reason</th>
<th>Real-name (n = 65) Modes</th>
<th>Anonymity (n = 65) Modes</th>
<th>Nickname (n = 66) Modes</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question type</td>
<td>% rank</td>
<td>% rank</td>
<td>% rank</td>
<td></td>
</tr>
<tr>
<td>Number of assessors already</td>
<td>82.81 1</td>
<td>69.35 1</td>
<td>65.63 1</td>
<td>70.41</td>
</tr>
<tr>
<td>given feedback on the item</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question item itself</td>
<td>21.88 5</td>
<td>25.81 5</td>
<td>25.00 5</td>
<td>23.47</td>
</tr>
<tr>
<td>Question submission time</td>
<td>32.81 4</td>
<td>30.65 4</td>
<td>31.25 4</td>
<td>30.61</td>
</tr>
<tr>
<td>Who the question-author is</td>
<td>15.63 6</td>
<td>4.84 6</td>
<td>21.88 6</td>
<td>13.78</td>
</tr>
<tr>
<td>The current average rating</td>
<td>59.38 3</td>
<td>48.39 3</td>
<td>65.63 1</td>
<td>56.12</td>
</tr>
</tbody>
</table>
Second, the students’ targeting behavior may be directed by various factors. In contrast to the possible adverse effects frequently discussed under identity revelation versus the merit of anonymous assessment literature, this study approached the issue from a different perspective—an investigation of peer assessors’ decision-making process in terms of which question they chose to assess. The non-significant results in the students’ targeting behavior implied that instructors may consider allowing students to have some freedom in this regard because of their differing motivations, interests, needs, and preferences (Ediger, 2006). Along the same humanist line of thought, instructors with access to online learning systems that do support anonymity or nickname use are advised not to arbitrarily impose anonymity and nickname modes since doing so may have adverse effects on interpersonal relationships (Yu & Wu, 2011).

Finally, as reflected in the students’ selections in the end-of-session checkbox question, of the six available options, on average, more than 50% of the participants saw “question type,” “number of assessors already giving feedback on the item,” and “the current average rating,” as the top three factors directing their online targeting behavior. Designers of technology-enhanced learning systems should include details of these features in user interfaces in order to facilitate peer-assessment activities.

Limitations of this study

The present study has some limitations. First, this study took place within intact classes, in which the fifth-grade participants were already familiar with each other prior to the experiment. As assessments may be affected by the existing friendships among students (Chan, 2010), who may guess who the real person is based on the nickname chosen, the findings may not be applicable to online open forums or communities where participants who may or may not know each other or to other age groups.

Second, use of a person’s real-name was set as the pre-treatment condition in this study. It is possible that questions from the same author may exhibit similar writing styles during pre-treatment and treatment conditions in such a way that student identity is revealed, even if nicknames are used or identities are later concealed.

Third, a checkbox question was used in this study to collect data on the participants’ accounts of their online targeting behavior. Although the six responses listed as possible answers were based on informal, non-participant observation of the participants’ selection behavior during the activities, other possible reasons may have been overlooked.

Fourth, the time allocated for interacting with the online system was fixed and limited to in-class use. In addition, the actual study was conducted in two instructional sessions, which may not have been long enough to allow for the effects implied by identity construction theory to manifest. The generalizability of this study to contexts without time and access restrictions and with extended implementation time should thus be exercised with caution.

Finally, the study involved Taiwanese elementary school pupils. Studies from the field of social communication, as well as those examining social perceptions and emotional perceptions, have found that students with different backgrounds have different cultural values (e.g., independent self vs. interdependent self) (Masuda, Ellsworth, Mesquita, Leu, Tanida & Van de Veerdonk, 2008) and exhibit different interaction patterns and behavior under identified and unidentified anonymity conditions in group situations (Chester & Gwynne, 1998; Hosack, 2004). Therefore, future studies involving different age groups with different cultural origins are thus needed to further extend the generalizability of the current work.

References


