Huang, C. Q., Wu, X. M., Wang, X. Z., He, T., Jiang, F., & Yu, J. H. (2021). Exploring the Relationships between Achievement Goals, Community Identification and Online Collaborative Reflection: A Deep Learning and Bayesian Approach. *Educational Technology & Society*, *24* (3), 210–223.

Exploring the Relationships between Achievement Goals, Community Identification and Online Collaborative Reflection: A Deep Learning and Bayesian Approach

Changqin Huang[1,2*], Xuemei Wu[2], Xizhe Wang[1], Tao He[2], Fan Jiang[1] and Jianhui Yu[1]
[1]Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University, Jinhua, China // [2]School of Information Technology in Education, South China Normal University, Guangzhou, China // cqhuang@zju.edu.cn // wuxuemei@m.scnu.edu.cn // xzwang@zjnu.edu.cn // tao.he2016@gmail.com // jiangfan1116@foxmail.com // jianhuiyu@126.com
[*]Corresponding author

ABSTRACT: Collaborative reflection (co-reflection) plays a vital role in collaborative knowledge construction and behavior shared regulation. Although the mixed effect of online co-reflection was reported in the literature, few studies have comprehensively examined both individual and group factors and their relationships that affect the co-reflection level. Therefore, this study explored the structural relationships between achievement goals (task-based, self-based, and other-based goals), online community identification, and co-reflection, which can consequently assist instructors in improving the related pedagogical strategies. To this end, 26813 posts on MOOC and college online learning platforms were gathered. Specifically, deep learning techniques were first used to train a classifier that classifies the large-scale co-reflection text automatically. The Bayesian method was then applied to disclose the structural relationships among achievement goals, community identification, and co-reflection. The results showed that the proposed classification algorithm achieved the best performance. Two best-fit models for characterizing the respective relationships between co-reflection and community identification as well as achievement goals were obtained using the Bayesian method. The results of the experiments on these two models demonstrated that both task-avoidance and other-avoidance goals were related directly to co-reflection, all task-approach, self-approach and other-approach goals were related indirectly to co-reflection, but self-avoidance goals had both a direct and an indirect relationship with co-reflection. The relationship between community identification and co-reflection was mediated by other-based goals. Some theoretical and practical implications were discussed for instructors and practitioners to build an online community.

Keywords: Deep learning, Bayesian network, Achievement goals, Co-reflection, Community identification

1. Introduction

Co-reflection refers to a process of collaborative critical thinking and knowledge construction, the activities of which are commonly affected by a combination of elements of individuals and groups (Kalk, Luik, & Taimalu, 2019). One way of supporting co-reflection is to use the tools provided by information and communication technology, such as blogs, e-portfolios, Facebook etc. In particular, these tools accommodate an open, flexible and diverse online learning community where students can reflect collaboratively on their thoughts, compared to expressing their thoughts in traditional ways (Yilmaz & Keser, 2016). Individuals would be motivated to read other peers' postings and comments, to develop a sense of community. In turn, more time is spent on their postings, which consequently may lead to an in-depth reflection (Clarà, Kelly, Mauri, & Danaher, 2017; Huang, Han, Li, Jong, & Tsai, 2019). However, researchers have assessed the level of online co-reflection with reporting mixed results. Some studies have shown that many students only describe or summarize what happened rather than critically think about it (Ozkan, 2019). Dalgarno, Reupert, and Bishop (2015) stated that some negative responses are given due to the lack of peer feedback, apparent resistance, and learning community engagement etc. However, few studies have investigated the antecedents and driving mechanism of online co-reflection, which can provide some theoretical and practical implications to motivate learners to be deeply engaged.

Previous researchers explored factors influencing co-reflection such as peer feedback and interactive behavior (Novakovich, 2016). However, individuals' participation in communities is for certain purposes, however learning motivation refers to some significant individual factors that guide and regulate individuals' behavior (Lim & Lim, 2020), which is the condition of intention to act (Chang, Hou, Wang, Cui, & Zhang, 2020). Therefore, there is a strong need to further investigate the factors influencing co-reflection from the perspective of motivation. Community identification is another crucial concept that facilitates members' participating, sharing, and knowledge constructing (Ergün & Avcı, 2018). It also plays a significant role in bridging the individual and group factors (Chang et al., 2020). Some studies have indicated that there may be different interactive relationships between community identification and achievement goals in a collaborative environment

(Chang et al., 2020; Thijs & Fleischmann, 2015). Therefore, this study was designed to explore the relationships among different achievement goals, community identification, and co-reflection in an online learning community.

In addition, the large-scale online discussion data and reflective writing provide valuable information to understand students' co-reflection, but also raise some problems of data analysis (Liu, Zhang, Wang, & Chen, 2017). Although features can be automatically captured from the data by machine learning methods, costly manual engineering is also required (Ullmann, 2019). Deep learning is a representation learning technique which can process the raw input to be suitable for the classification of feature engineering, and it has been recognized as the most advanced solution to performing tasks in data mining related to classification (LeCun, Bengio, & Hinton, 2015). However, few works have applied deep learning techniques to analyze reflective texts (Chen, Xie, Zou, & Hwang, 2020).

For this research, the deep learning technique and Bayesian method are applied to make the automatic prediction of online co-reflection levels, as well as discover the relationships between achievement goals, community identification and co-reflection. Specifically, two research questions (RQ) are proposed in this study:

RQ1: To what extent can the deep learning technique accurately classify the level of co-reflection of each student?
RQ2: What are the relationships between achievement goals, community identification and co-reflection?

2. Theoretical background

2.1. Achievement goal theory

Achievement goal theory is a predominant theoretical framework of achievement motivation to interpret different qualities of individual learning and well-being, particularly in educational contexts (Urdan & Kaplan, 2020). Various models of the achievement goal theory have been proposed to conceptualize students' motivational orientations to understand students' motivational beliefs, their causes and effects (Elliot, Murayama, & Pekrun, 2011; Elliott & Dweck, 1988). Existing studies emphasized that learning motivation and achievement goals provided an essential foundation for reflection and meaning construction (Anderman, 2010; Tikhomirova & Kochetkov, 2018). Some researchers indicated that learners might have diverse goal-oriented motivation mechanisms in different contexts, e.g., individual versus collaborative learning environments (Lim & Lim, 2020). Thijs and Fleischmann (2015) pointed out that achievement goals depended on individuals' perception of relatedness to others. Therefore, this research will further explore the driving mechanism of achievement goals on co-reflection in a collaborative learning community.

2.2. Social identity theory

Social identity theory (SIT) provides an essential theoretical background for community identification and member behavior, which indicates that group members establish their identity in a community by viewing themselves as a part of that, and generating an emotional attachment to the group or community (Tajfel, 1978). It should be noted that social identification involves not only perceived self-categorization, but also the evaluative and affective states with the social group, and this identification with the group allows members to modify their thoughts and behaviors (Qu & Lee, 2011). Chang et al. (2020) found that community identification significantly mediated the relationship between motivation and members' community behavior. Additionally, Bowskill (2017) pointed out that inducing a sense of group identity can motivate self-evaluation and critical thinking engagement within a technology-supported learning community. Therefore, the learners' sense of identity with the group might be an important factor influencing co-reflection in this study.

Further, considering that co-reflection is a process of knowledge co-construction, including individual and group cognition, it is necessary to comprehensively investigate the essential individual and group factors that affect it. Grounded on achievement goal theory and community identification theory and the related research, this study mainly focuses on two pivotal factors, achievement goals and the learners' sense of identity with the group and reveals their driving mechanism for online co-reflection.

3. Literature review

3.1. Co-reflection

Co-reflection is a process of collaborative critical thinking involving cognitive and affective interactions between two or more individuals who explore their experiences to reach new intersubjective understandings and appreciations (Yukawa, 2006). This definition of co-reflection brings new perspectives and considerations from the dialogue with others who might see situations differently, challenge assumptions, or ask significant questions (Krutka, Bergman, Flores, Mason, & Jack, 2014). These arguments are consistent with those by Vygotsky (1978) who assumed that cognition is a process of social interaction with each other. In this study, we also believe that co-reflection would be deepened when engaged in communion with peers who could push each other beyond description to thoughtful reconsideration (Krutka et al., 2014). However, existing studies mainly explored platforms or strategies that support co-reflection. Kalk, Luik, and Taimalu (2019) reported that the reflection level can be predicted by the characteristics of students, blog groups and blogging. But the essential factors that affect the level of co-reflection and its driving mechanism are still lacking.

3.2. Achievement goals

Achievement goals are the integrated systems, theories, or schemas, that incorporate conceptions of ability, perceptions of the self and features of self-consciousness, definitions of success in specific achievement contexts, and affective and behavioural responses (Urdan & Kaplan, 2020). Recently, the latest achievement goal theory model proposed by Elliot, Murayama, and Pekrun (2011) offers a six-component model, which includes task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance. And all of them are distinguished by task, self, and other three competence evaluation standards. Elliot and Thrash (2001) remarked that six possible types of achievement goals as the basis for evaluation have many benefits, that is, it explicitly accounts for both the energization and direction of competence-based behaviour, and provides a more specific definition of the achievement goal construct. Also, it affords greater conceptual flexibility in that any combination of reason and goal may be considered. Therefore, this model was adopted as one conceptual framework in the present study.

Additionally, although several studies have been conducted to explore the relationship between achievement goals and reflection (Mercier, 2017), a consensus was not reached about the effects of different achievement goals on reflection (Urdan & Kaplan, 2020). Moreover, studies on the relationship between achievement goals and reflection mainly focus on individual reflection (Collin & Karsenti, 2011). Thus, it is meaningful to conceptualize the effects and driving mechanism of different achievement goal orientations on co-reflection in an online learning community.

3.3. Community identification

According to social identity theory, community identification refers to the degree to which individuals feel a sense of belonging to the community (Tajfel, 1978). Feeling like part of the group in a community is considered a critical factor for a successful online community building (Qu & Lee, 2011), and members with a high level of identification can reduce their stress, enhance their self-esteem and be motivated to modify their thoughts and behaviors according to the group's common values and interests (Chiu, Huang, Cheng, & Sun, 2015). Recently, attention was given to this potential pathway that links community identification and community participation, members' knowledge sharing and construction (Yilmaz, 2016). Thus, exploring the relationship between different achievement goals and community performance will help us better understand the individual's behavior.

3.4. The relationships between achievement goals, community identification and co-reflection

Previous studies have explored the relationship between achievement goals and co-reflection, showing that there are different direct and indirect relationships between them. Mercier (2017) found that although learning and performance goals displayed no differences in outcome measures, groups with the former goal showed more reflection and explanations than groups with the later goal during the task. Lau, Liem, and Nie (2008) reported that task-approach and task-avoidance goals have both a direct and an indirect effect on deep learning, and the relationship between the two of them and deep learning is mediated by classroom attentiveness and group participation. However, group participation mediated the relationship between the performance-approach goal and deep learning. Chang et al. (2020) pointed out community identification significantly mediated the relationship between motivation and social loafing. Therefore, it can be inferred that the relationship between achievement goals and co-reflection may be mediated by community identification, and different goal orientations may have different indirect or direct relations to co-reflection.

Conversely, some studies indicated that the other different potential path exists between achievement goal, community identification and co-reflection. For example, Zumbrunn, McKim, Buhs, and Hawley (2014) found that expectancy (one of the important motivational constructs) significantly mediated the relationship between sense of belonging (a construct like community identification) and achievement. But task values failed to mediate the relations. Further, Won, Wolters, and Mueller (2018) examined the relationships between sense of belonging, achievement goals and self-regulated learning, reporting that only mastery goals mediated the relationship between the sense of belonging and metacognitive. This implies that students' identification affects the achievement goals or the reasons or purposes they used in the task, which in turn impact their academic effort and engagement (Won et al., 2018). The self-determination theory (SDT) can provide some supportive evidence for this, which underlined the need for relatedness to others plays a critical role in students' motivation and performance (Van den Broeck, Ferris, Chang, & Rosen, 2016). Therefore, different goal orientations will be affected by community identification with varying degrees.

Taken together, there may be two different potential relationships between achievement goals, community identification and co-reflection. However, further two important gaps need to be noted and filled. First, although existing research investigated the relationship between achievement goals and reflection, the accurate relationships between different goal orientations, community identification and co-reflection are still unknown. Therefore, this study explored the relationship between co-reflection and achievement goals based on the six-factor achievement goals model. Second, prior studies mostly proposed a hypothetical model and used the structural equation modelling method to further verify the fitting effect, which is theory-driven. Instead, this study attempts to mine the relationships between different achievement goals, community identification and co-reflection using the Bayesian method from a data-driven perspective.

3.5. Deep learning for educational applications

Deep learning has a multilayer network structure and has a strong power to learn discriminative information from examples, patterns or events (Waheed et al., 2020). Many applications, such as learning performance prediction, learning recommendation, intelligent learning tool and system development, have been explored based on various methods (Hwang, Sung, Chang, & Huang, 2020; Hwang, Xie, Wah, & Gašević, 2020; Wang, Mei, Huang, Han, & Huang, 2021; Zhou, Huang, Hu, Zhu, & Tang, 2018). The most commonly used method is text classification in educational data mining (Chen, Xie, & Hwang, 2020; LeCun et al., 2015). Ullmann (2019) concluded that there are three approaches (machine learning-based, dictionary-based and rule-based) for reflective text analysis. However, all of these have their limitations (e.g., costly manual feature engineering, time-consuming etc.). Deep learning has great potential for educational data mining, especially in text classification (Young, Hazarika, Poria, & Cambria, 2018). Therefore, deep learning is conducted for co-reflection text classification in this study.

4. Methodology

4.1. Research design

To answer the two research questions, this study consists of four stages, as depicted in Figure 1. Specifically, students' online co-reflection text data and the questionnaire data of achievement goals and community identification was collected. Furthermore, the text and questionnaire data were further preprocessed to ensure validity. For RQ1, this study adopted the techniques of BERT and LSTM to classify reflective texts to identify students' co-reflection level, then the performance of the classification model was evaluated. For RQ2, the trained classification model was used to identify each student's level of co-reflection, and the Bayesian method was then integrated to explore the relationship between the three factors (online co-reflection, achievement goal, community identification).

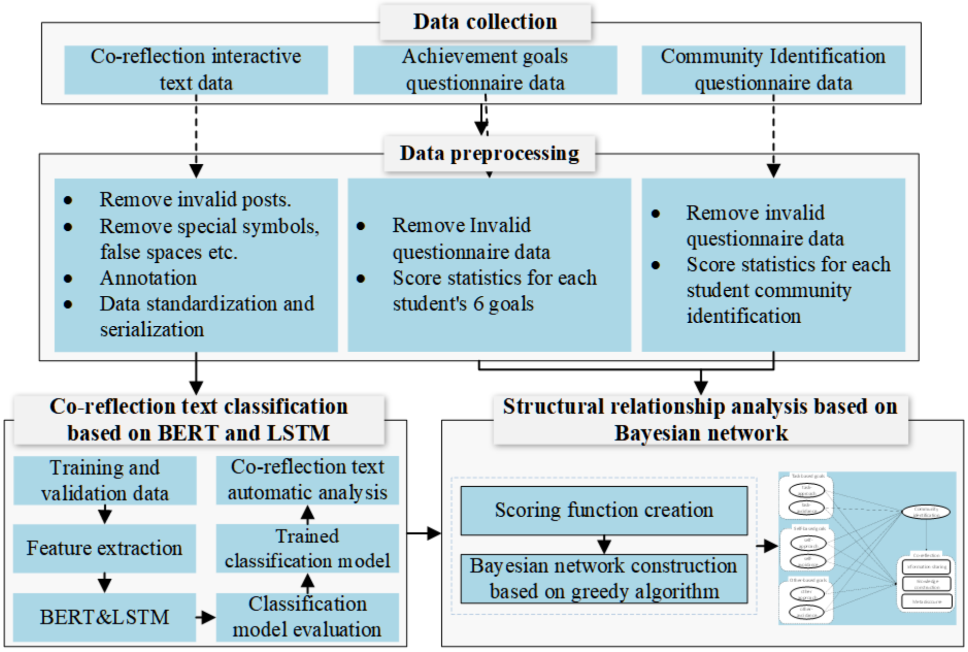

Figure 1. The research design of this study

4.2. Data collection and preprocess

The co-reflection text data was collected from the three online courses on educational technology, with each course being offered for almost five months. During the course, learners participated in online co-reflection activities in a similar way that each discussion and reflection began after the topic was posted. From a total of 26813 original posts collected, 16890 posts were determined as the dataset after removing the invalid data. To reduce noise in the dataset, all duplicate posts and special symbols such as punctuation marks, false spaces and emoticons were removed according to Liu's et al. (2017) recommendation.

The data for the other two variables under consideration were gathered through a Chinese survey website. In the survey, a total of 115 undergraduate and graduate students who participated in an average of 12 to19 online co-reflection activities in the two courses were invited to complete the questionnaires voluntarily. In the first class of the course, students were invited to fill in the achievement goal questionnaires adopted from Elliot et al. (2011) which comprised 18 measurement items. At the end of the course, students filled in the community identification questionnaires adopted from Chang et al. (2020) comprising four measurement items. All the items were measured on a five-point Likert scale ranging from "not true of me (1)" to "extremely true of me (5)." Finally, a total of 95 valid responses were obtained. The Cronbach's alpha of achievement goal and community identification were 0.916 and 0.898 respectively.

To train an efficient classifier, a structural dataset with label information was constructed for training and testing the classification model. To do this, the unit of analysis was defined as a complete dialogue with the same peer on each topic, also called an episode (Mercier, 2017). Each unit of the analysis was coded by two coders according to Lei and Chan's (2018) coding scheme. Specifically, the scheme consists of nine reflection levels (see Table 1), in which 1 to 3, 4 to 6, and 7 to 9 are reconsidered as low, middle, and high three levels of sharing of information, knowledge construction, and metadiscourse, respectively. In this study, each analysis unit was marked as one of three levels 1, 2, and 3 respectively, with a unit that does not belong to any of the nine categories marked as 0. Discussions and revisions were undertaken among the research team members until consensus was reached on each post. Finally, Cohen's Kappa was computed as 0.878 ($p < .01$), which indicates a high level of agreement between the coders.

Table 1. The coding scheme of co-reflection levels

Categories	Description	Exemplar excerpts from co-reflective logs
1. Listing and Copying	Lists notes without explanations; copies information from or repeat other's notes in a very close way	Share an article "coupled teacher" or "double loss," see the link below.
2.Brief Summary	Summarizes a few notes shortly and often incompletely	By summarizing the views in the two articles, the principles of CAI courseware design are as follows: 1. Educational principles 2. The principle of control...
3.Interpretation or Elaboration	Interprets others' notes on information with different wording or extends information by examples or evidence	The previous students mentioned many professional tools, and almost gave a detailed overview of…, I still tend to recommend the two most commonly used tools, blog and WeChat...
4.Question-Based Discussion	Sees the discussion as question-based and a deepening process of seeking answers to questions	In my view, the focus of educational technology is technology, I think educational technology is …
5.Constructive Use of Information	Uses information, either from experts, books, the Internet, or other related courses, life experience, etc. to justify or deepen ideas	Once we visited the teacher, he suggested that we should pack the knowledge of each chapter and put the packed knowledge in different boxes...I think the process of finding and marking boxes is the process of building knowledge scaffolding, because...
6.Intertwined Question Explanation	Keeps asking related questions, showing doubt or seeking clarification; responses and explanations are intertwined progressively in the discussion	Can the cultivation of innovation ability be reflected through their group discussion process? ... For example, encourage them to innovate in the display of the discussion results, etc.
7.Meta-Cognition	Reflects on what the class does not know; realizes high points in the discussion; self-defines goals and tasks for exploration	Our current progress is to learn some artificial intelligence knowledge, I think the purpose is to be able to understand the relevant papers. The first step of the next plan is to improve academic literacy
8.Meta-Theory	Focuses on theories while developing the discourse; uses theories/conjectures to explain the phenomena, even making attempts to create new theories	…When I mentioned how to balance curriculum planning, I thought of Cuba's thought-provoking point...It must be explained that the emergence of information technology has raised the issue of curriculum design... Therefore, education and technology themselves are also a pair of balanced propositions.
9. Meta-Conversation	Focuses on examining what the discourse is about, especially reflecting on discourse goals; adopts a "we" perspective to assume collective responsibility for advancing knowledge; tackles difficult/important issues which may be neglected by the community	Yes, there is a discussion that can produce a collision of ideas...So the purpose of the mutual evaluation is designed to urge the group members to participate in the group discussion more seriously.
10.Other	Some posts include greetings, thanks, simple compliments, etc.	Very good! Thank you! Morning! etc.

4.3. Co-reflection text classification based on BERT and LSTM

Previous studies have shown that it is difficult for students to achieve a deep level of reflection in a short time, and the quality of the reflection is related to the mastery of knowledge (Granberg, 2010; Van den Kieboom, 2013). Differing from the existing classification models, the long short-term memory (LSTM) model that can capture long-term dependencies (Yu, Si, Hu, & Zhang, 2019) was therefore employed to obtain the time series information of the reflective text. The BERT (Bidirectional Encoder Representations from Transformers) pre-training model performs the best on language understanding and text extraction (Devlin, Chang, Lee, & Toutanova, 2018). Therefore, BERT and LSTM were integrated to classify co-reflection levels from a large-scale dataset. The overall BERT and LSTM architecture of our classification model can be seen in Figure 2. A total of 10572 labelled posts were used as the training dataset and the reflective text data of each student was arranged in chronological order of reflection topics (Topic1, Topic2, …Topic *m*). Each post was segmented and vectorized based on the jieba library as well as BERT's pre-training. Vectorized and positioned co-reflection text information was obtained and used as input to the BERT model for fine-tuning. In this way, a serialized vector of the co-reflection text from each topic was obtained (C_1, C_2, …C_m). It was later taken as the input of the text classifier based on the LSTM model. Finally, a fully connected (FC) layer and Softmax function were used to classify the output vector of LSTM and generate the final prediction result.

For training, ten-fold cross-validation was used for each algorithm as well as the metrics of accuracy, precision, recall, and F1 (the harmonic mean of precision and recall) which are commonly used to evaluate the performance of text classification tasks and measure the proportion of correct predictions from different perspectives (Hew, Hu, Qiao, & Tang, 2020). Therefore, these metrics were employed to measure the performance of the classification model in this study. Other pre-training models (e.g., Word2vec), serialization analysis methods (e.g., historical average (HA)), and keywords methods (e.g., TF-IDF) were also implemented for the classification task in the present research. After the training process, the text classification model was used to automatically classify the rest of the texts into different levels of co-reflection. The resulting levels will be used for structural relationship analysis presented in the next section.

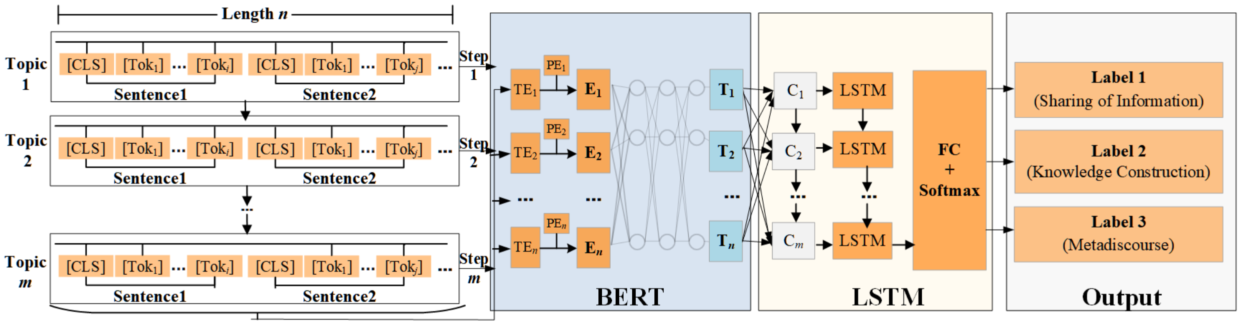

Figure 2. Co-reflection level prediction model integrating BERT and LSTM
Note. *m* = the total topic length; *n* = the number of words in each topic; TE = text embedding, and PE = the position embedding of text.

4.4. Structural relationship analysis based on Bayesian networks

The Bayesian method of directed association mining was used to explore the relationships between achievement goals, community identification and co-reflection. Traditional confirmatory data analytic procedures (e.g., path analysis, structural equation modelling) follow a "frequentist" approach to test a network of effects in a model. Such approaches generally do not fit well with the prescribed model, which may lead to improving the model fit by practice with inherent problems (Hagger & Hamilton, 2018). Instead, the Bayesian approach assumes that model parameters have inherent uncertainty that is represented by a distribution. As a powerful tool that infers uncertain association relationships, the Bayesian approach has been widely used to automatically mine the associations and causal relationships between factors (Heckerman, 1997). It is well suited for exploring the relationships between achievement goal orientations, community identification and co-reflection in this study. However, some researchers recommend that caution should be taken and more prior information ought to be considered when using the Bayesian approach (Hagger & Hamilton, 2018; Meyer & Xu, 2007). Therefore, in this study, the theoretical prior knowledge is comprehensively considered with the Bayesian structure analysis.

To determine the optimal structural relationship between the three factors, eight nodes in a Bayesian network were constructed firstly, and then the most appropriate network structure was selected from the existing datasets. The distinctions between six orientations of goals have been validated by multiple studies (Elliot et al., 2011). In this study, the internal logical relationships between the six types of goals were therefore regarded as controlled. Achievement goals, community identification, and co-reflection were then set to the first, second, and third levels of the network, respectively. After that, an optimal model would be determined based on the existing data. Then, the positions of achievement goal and community identification in the network were swapped, and the same operation was repeated for others.

To improve the model further, the scoring function method was implemented to evaluate the degree of fitting between the Bayesian network and training dataset. As such, whether to add, remove or adjust the directions of the edges of the Bayesian network was determined by looking at the changes of the score. Note that the nodes represent the variables and the edges indicate the relationship between two variables in the Bayesian network. That is, changes to the edges are equivalent to exploring possible relationships between the variable. Specifically, in Figures 3 and 4, the BDeu ("BD" for Bayesian Dirichlet, "e" for likelihood-equivalence, "u" for uniform joint distribution) score, K2 score and Bic score measure the degree of fit. The larger the value, the better the model fits (Carvalho, 2009). In addition, a greedy algorithm was used to identify a stable relationship structure by continually updating until the score function value remains unchanged. In this, a relatively stable network relationship structure can finally be obtained.

5. Results

5.1. Co-reflection text classification results for RQ1

Table 2 lists the performance results of the classification models with different algorithms. The classification model which integrates BERT and LSTM performed better than the other models. Specifically, the pre-training model based on BERT (e.g., BERT & LSTM, BERT & HA) performed better than Word2Vec (e.g., Word2Vec & LSTM, Word2Vec & HA), and TF-IDF performed the worst. Furthermore, the algorithms that integrated the pre-training model and the serialization model (e.g., BERT & LSTM, Word2Vec & LSTM) performed better than those without the serialization model (e.g., BERT & HA, Word2Vec & HA). Again, the algorithms that combine BERT and LSTM performed the best. However, the results also revealed that these algorithms were less effective than human judgments. In particular, the performance of our algorithm in terms of precision, recall, accuracy, and F1 was an average of 3.7% lower than human judgments. According to the literature, the error was acceptable within 10% (Ullmann, 2019). Generally, the classification model of integrating BERT and LSTM demonstrated reasonably good performance.

Table 2. Text classification model results

	Precision	Recall	Accuracy	F1
Human	81.25%	78.00%	78.95%	79.63%
TF-IDF	58.33%	58.33%	57.89%	58.33%
Word2Vec & HA	62.50%	63.83%	63.16%	63.16%
Word2Vec & LSTM	66.67%	69.57%	68.42%	68.12%
BERT & HA	64.58%	70.45%	68.42%	67.52%
BERT & LSTM	75.00%	76.60%	75.79%	75.80%

5.2. Structural relationship results for RQ2

Two models were chosen (see Figures 5 and 6) through multiple rounds of evaluation and selection. Figure 3 and Figure 4 show the trend of the BDeu, K2, and Bic scores of the two models respectively as the number of edges of the model decreased. As shown in Figure 3, as the number of edges in the model decreases, the score of BDeu becomes larger. But BDeu and Bic scores tend to be flat when five edges in this model have been removed. Continuously, an obvious downward trend of the K2 score was observed when six edges in the model have been removed. This indicates the fit degree between the model and data is relatively higher without the need to provide more information. Taken together, an approximate optimal model (model 1) was obtained. In the same way, model 2 was also obtained.

For model 1, all eight variables were selected from the competing admissible models. The conditional probability of each variable was computed according to the standardized values (0, 1, 2) converted from the original scores of achievement goals and community identification variables. It measures the degree of the links between different variables. According to the selected model 1, this shows that community identification mediated the relationship between achievement goals and co-reflection. Specifically, both the task-avoidance goal and other-avoidance goal have direct relations to co-reflection, while the three goals of the task-approach, self-approach and other-approach have an indirect link to co-reflection. Interestingly, the self-avoidance goal has both a direct and an indirect path to co-reflection.

For model 2, eight variables were similarly retained after the selection. Specifically, this admissible model showed that community identification has only indirect connections to co-reflection, and the relation is mediated by other-based goals. The task-avoidance goal and self-avoidance goal only has a direct relation to co-reflection, respectively, whereas the goals of task-approach and self-approach have no direct relation to co-reflection.

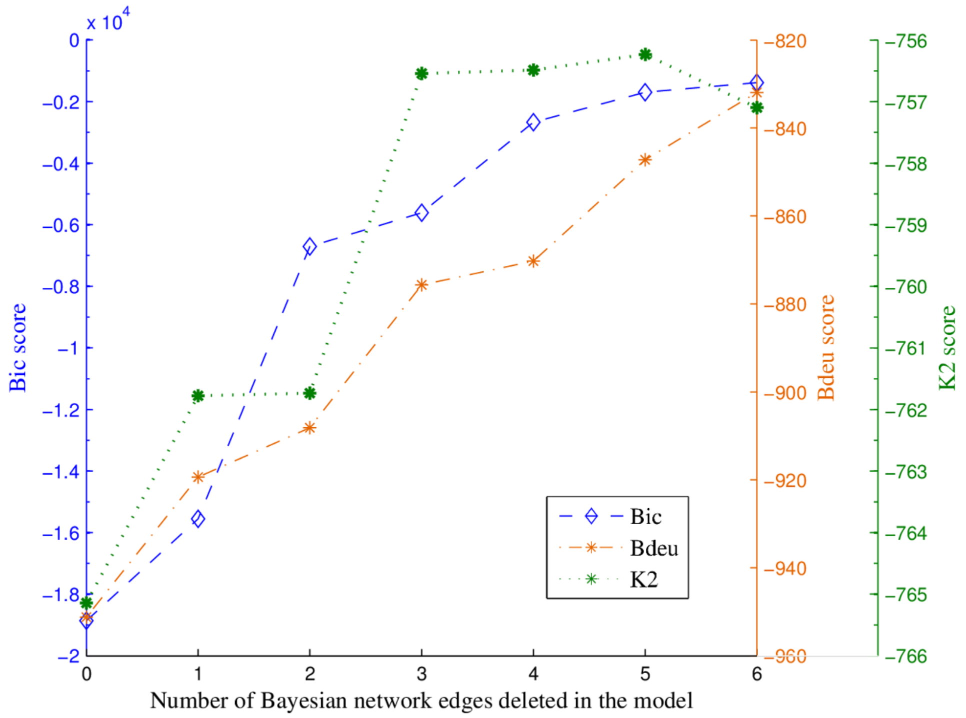

Figure 3. Bdeu, K2, Bic score in model 1

Note. 0-6 in the figure in the x-axis is the number of Bayesian network edges deleted in the model. The y-axis is the score.

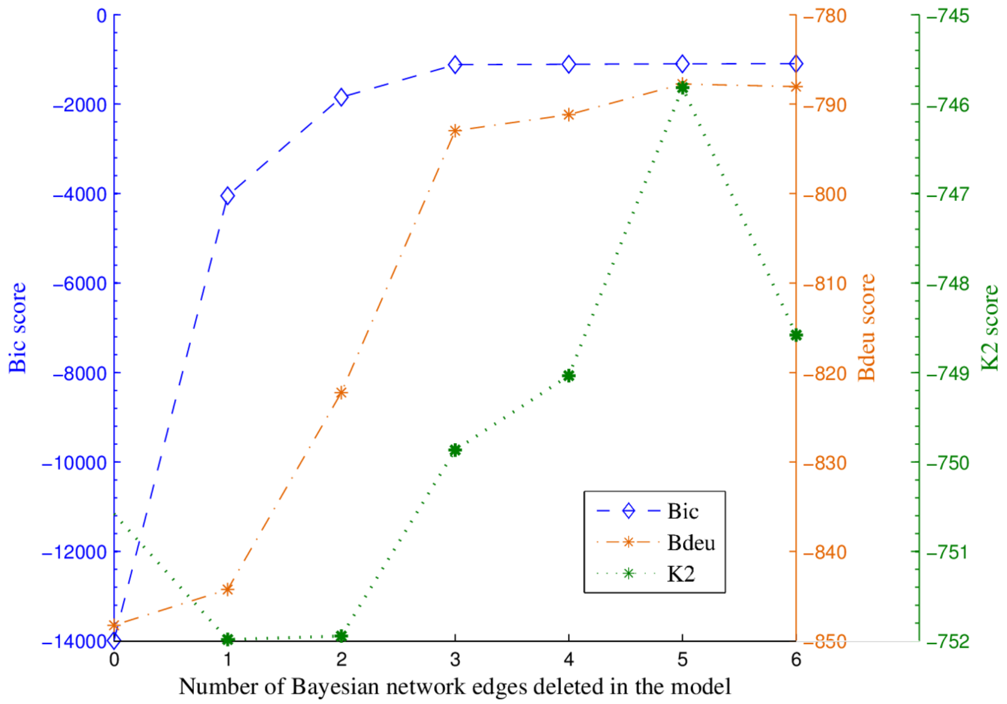

Figure 4. Bdeu, K2, Bic score in model 2

Note. 0-6 in the figure in the x-axis is the number of Bayesian network edges deleted in the model. The y-axis is the score.

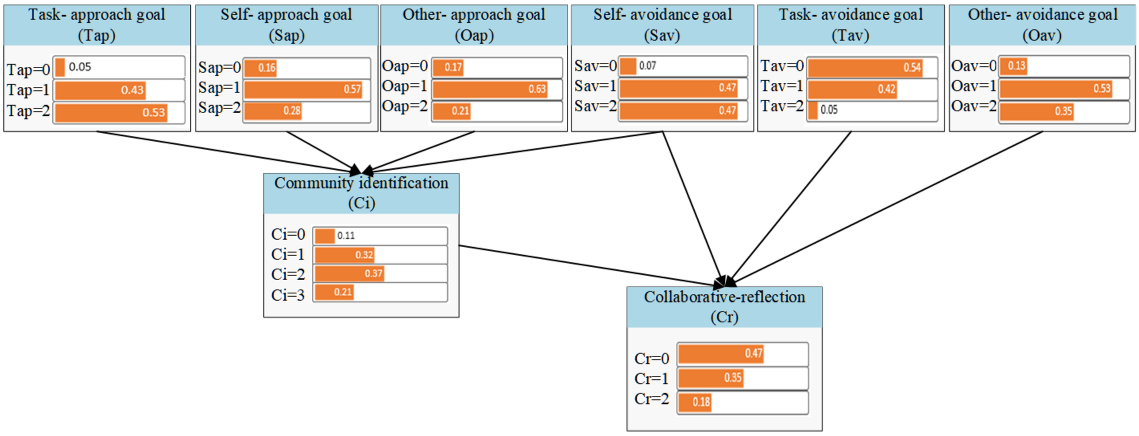

Figure 5. The structural relationship of model 1

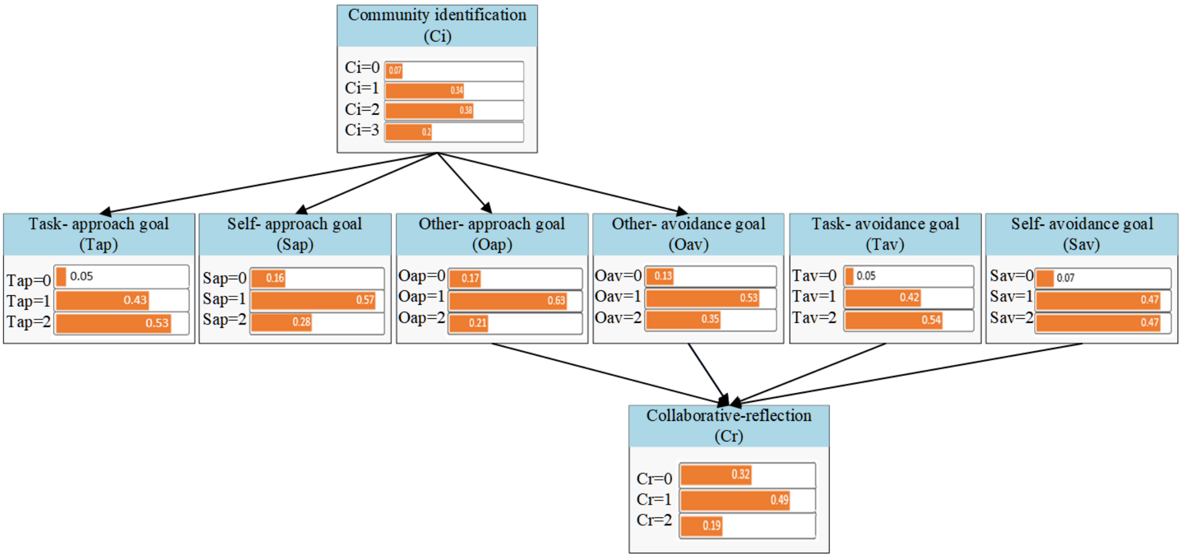

Figure 6. The structural relationship of model 2

6. Discussion

To analyze the online large-scale interactive text data about students' co-reflection information, the present study combined the BERT pre-training model and LSTM into an integrated classifier that performed better than the baseline models. On the one hand, the BERT pre-training model uses the mask method and has migration capabilities (Devlin et al., 2018) which can quickly and precisely understand the feature of reflective text language in this study. On the other hand, the BERT model with the embedded attention mechanism, which is not limited to the length of the text sequence, can improve the accuracy of the classification model compared to conventional methods (González-Carvajal & Garrido-Merchán, 2020). In addition, the time series feature of the reflective text is captured based on LSTM, which is in line with the actual development of the learner's level of reflection and accords with reflection as the essential feature of the internal cognitive process (Granberg, 2010). Therefore, the integrated classification model can more accurately identify the co-reflection level and this model also confirmed the advantages of using deep learning techniques for educational data mining, especially for text classification tasks (Young et al., 2018).

Along with the two best-fit models found in this study, the different potential relationships between achievement goals, community identification and co-reflection were indicated. For task-based goals, task-approach goals were not directly related to co-reflection, and community identification mediated the links of the task-approach goal and co-reflection. This is not completely consistent with the existing conclusion that mastery goals were both directly and indirectly related to deep learning strategies and outcomes (Heo, Anwar, & Menekse, 2018; Lim & Lim, 2020). There are a few possible explanations for this inconsistency. First, the students were required to participate in co-reflection activities that interacted with others, and this may push them to reach intersubjective understandings. That is, in this process, they would have sense of community, which in turn affected their co-reflection further. This conforms to previous studies (Lau et al., 2008). Another possible explanation is that the community identification perceived by the students may increase their task value (David, 2014), which in turn promoted their participation in co-reflection. This assumption was also made by Zumbrunn et al. (2014), but further investigation is still needed. Furthermore, for task-avoidance goals, positive or negative relationships between task-avoidance goals and help-seeking behaviors have been described in previous studies. Based on the present results, relatedness to others, however, may not be the main psychological need that motivated these students to work hard and participated in co-reflection (Van den Broeck et al., 2016). According to Elliot et al. (2011), students with task-avoidance goals mainly attained satisfaction by completing challenging tasks. Therefore, further investigations on considering the factor of the task value may be more helpful to understand the relationships between task-based goals, community identification, and co-reflection.

For other-based goals, the result indicated that the depth of co-reflection for the students was mainly affected by their perceived community identification. It may regulate their learning strategies and goals for participating in co-reflection. This is not entirely consistent with the existing conclusion (Lau et al., 2008; Won et al., 2018). But as for the psychological need and competence evaluation criteria of the task-approach goal, the results implied in the present study is consistent with existing findings (Elliot et al., 2011; Van den Broeck et al., 2016). Besides, community identification did not mediate the relationship between the other-avoidance goal and co-reflection. According to Payne, Youngcourt, and Beaubien (2007), students with other-avoidance goals had low help-seeking behavior and a sense of efficacy. They may be afraid of showing incompetence in front of their peers, with an attitude of resistance and avoidance to the community. Therefore, community identification would not mediate its relations to co-reflection unless the community identification was enough to allow them to regulate their own goals, and the performance of co-reflection could be promoted. Overall, students with other-based goals would regulate their goals through community identification in a way that affected their performance on co-reflection. According to SDT, different from the task-based goal, students with other-based goals may mainly take relatedness to others as their main psychological needs (Van den Broeck et al., 2016). But they may have different ways of behavioral regulation. For students with the other-avoidance goal, autonomous regulation and controlled regulation were dominant, while students with the other-approach goal possibly had controlled regulation (Deci & Ryan, 2012).

For self-based goals, students with self-approach goals had higher internal motivation and help-seeking behaviors (Elliot et al., 2011). Therefore, if they received help from their group, they may have a higher level of a sense of belonging. This could encourage them to share their knowledge and promote co-reflection. This fact was in line with the principle of cooperative reciprocity and the claims of SIT (Chiu et al., 2015). Unlike the self-avoidance goal, students with the self-approach goal, however, can regulate their own goals due to their perceived community identification, which was inconsistent with the finding of Elliot (Elliot et al., 2011). We inferred that the self-approach goal followed the competence evaluation criteria of self-improvement, but one would have more satisfaction and efficacy, tending to be in line with the group's common values and interests after an individual's goals were accepted by their group (Chiu et al., 2015). Moreover, the self-avoidance goal had a direct and an indirect link to co-reflection, but it was not affected by community identification. This conformed with Elliot's et al. (2011) view. The self-avoidance goal was based on self-improvement as the competence evaluation standard, and they feared performing worse than they had performed before. Therefore, their goals would not change due to the community identification they felt.

7. Conclusion

This research has comprehensively examined the factors that affect online co-reflection. In particular, the different relationships between achievement goals, community identification, and co-reflection were revealed using deep learning techniques and Bayesian methods. This work has made the following contributions. First, the present study is one of the few works that applies deep learning techniques to classify reflective texts to identify the learner's co-reflection level, which provides a methodological foundation for the construction of a platform which automatically monitors learners' co-reflection level. Second, this study has further validated the six-factor achievement goal framework by demonstrating the significance of the achievement goal theory in the context of online collaborative learning. Third, some practical implications can be provided for online community builders and instructors according to the driving mechanism of co-reflection found in this research. Specifically, to promote learners' in-depth co-reflection, practitioners should comprehensively consider learners' achievement goal orientations and community identification for providing the corresponding guidance.

The study also has several limitations. First, the data were collected using different methods. This may lead to deviations among different evaluation standards, affecting the accuracy of the results to some degree. Second, this research does not consider the causal relationship between different factors, using an exploratory attempt instead. Third, this research mainly focuses on theoretical and methodological exploration, but lacks practical educational applications.

There are some possible directions for extending this research. First, multi-modal and longitudinally serialization data should be collected to examine the relationships between achievement goals, online community identification and collaborative reflection more deeply. Second, the state-of-the-art language understanding and feature extraction methods like RoBERTa, ALBERT, and XLNet can be considered in further research. In addition, implementing educational applications and evaluating the effects according to the findings of this study are promising, such as developing a co-reflection platform or a personalized feedback system (Xie, Chu, Hwang, & Wang, 2019), exploring the integrating of the co-reflection platform and teaching (Zou, Xie, Wang, & Kwan, 2020), and investigating the feedback of teachers and students (Hwang, Yang, & Wang, 2013).

Acknowledgement

This work was supported by the Humanities and Social Sciences Planning Fund of the Ministry of Education, China (No. 18YJA880027).

References

Anderman, E. M. (2010). Reflections on Wittrock's generative model of learning: A Motivation perspective. *Educational Psychologist*, *45*(1), 55-60.

Bowskill, N. (2017). Sharedthinking: A Social identity approach to critical thinking. *Journal of Pedagogic Development*, *7*(2), 37-46.

Carvalho, A. M. (2009). Scoring functions for learning Bayesian networks. *INESC-ID Technical Reports*, *12*. Retrieved from http://www.lx.it.pt/~asmc/pub/talks/09-TA/ta_pres.pdf

Chang, Y., Hou, R. J., Wang, K., Cui, A. P., & Zhang, C. B. (2020). Effects of intrinsic and extrinsic motivation on social loafing in online travel communities. *Computers in Human Behavior*, *109*, 106360. doi:10.1016/j.chb.2020.106360

Chen, X., Xie, H., & Hwang, G. J. (2020). A Multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers & Education: Artificial Intelligence*, *1*, 100005. doi:10.1016/j.caeai.2020.100005

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100002. doi:10.1016/j.caeai.2020.100002

Chiu, C. M., Huang, H. Y., Cheng, H. L., & Sun, P. C. (2015). Understanding online community citizenship behaviors through social support and social identity. *International Journal of Information Management*, *35*(4), 504-519.

Clarà, M., Kelly, N., Mauri, T., & Danaher, P. A. (2017). Can massive communities of teachers facilitate collaborative reflection? Fractal design as a possible answer. *Asia-Pacific Journal of Teacher Education*, *45*(1), 86-98.

Collin, S., & Karsenti, T. (2011). The Collective dimension of reflective practice: The How and why. *Reflective Practice*, *12*(4), 569-581.

Dalgarno, B., Reupert, A., & Bishop, A. (2015). Blogging while on professional placement: Explaining the diversity in student attitudes and engagement. *Technology, Pedagogy and Education*, *24*(2), 189-209.

David, A. P. (2014). Analysis of the separation of task-based and self-based achievement goals in a Philippine sample. *Psychological Studies*, *59*(4), 365-373.

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. V. Lange, A. W. Kruglanski & E. T. Higgins (Eds.), *Handbook of Theories of Social Psychology* (Vol. 1, pp. 416-437). Thousand Oaks, CA: Sage.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. Retrieved from https://arxiv.org/abs/1810.04805

Elliot, A. J., Murayama, K., & Pekrun, R. (2011). A 3×2 achievement goal model. *Journal of Educational Psychology*, *103*(3), 632-648.

Elliot, A. J., & Thrash, T. M. (2001). Achievement goals and the hierarchical model of achievement motivation. *Educational Psychology Review*, *13*(2), 139-156.

Elliott, E. S., & Dweck, C. S. (1988). Goals: An Approach to motivation and achievement. *Journal of Personality and Social Psychology*, *54*(1), 5–12.

Ergün, E., & Avcı, Ü. (2018). Knowledge sharing self-efficacy, motivation and sense of community as predictors of knowledge receiving and giving behaviors. *Educational Technology & Society*, *21*(3), 60-73.

González-Carvajal, S., & Garrido-Merchán, E. C. (2020). Comparing BERT against traditional machine learning text classification. Retrieved from https://arxiv.org/abs/2005.13012

Granberg, C. (2010). Social software for reflective dialogue: Questions about reflection and dialogue in student teachers' blogs. *Technology, Pedagogy and Education*, *19*(3), 345-360.

Hagger, M. S., & Hamilton, K. (2018). Motivational predictors of students' participation in out-of-school learning activities and academic attainment in science: An Application of the trans-contextual model using Bayesian path analysis. *Learning and Individual Differences*, *67*, 232-244. doi:10.1016/j.lindif.2018.09.002

Heckerman, D. (1997). Bayesian networks for data mining. *Data Mining and Knowledge Discovery*, *1*(1), 79-119.

Heo, D., Anwar, S., & Menekse, M. (2018). The Relationship between engineering students' achievement goals, reflection behaviors, and learning outcomes. *International Journal of Engineering Education*, *34*(5), 1634-1643.

Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2020). What predicts student satisfaction with MOOCs: A Gradient boosting trees supervised machine learning and sentiment analysis approach. *Computers & Education*, *145*, 103724. doi:10.1016/j.compedu.2019.103724

Huang, C. Q., Han, Z. M., Li, M. X., Jong, M. S. Y., & Tsai, C. C. (2019). Investigating students' interaction patterns and dynamic learning sentiments in online discussions. *Computers & Education*, *140*, 103589. doi:10.1016/j.compedu.2019.05.015

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A Meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, *136*(3), 422-449.

Hwang, G. J., Sung, H. Y., Chang, S. C., & Huang, X. C. (2020). A Fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. *Computers & Education: Artificial Intelligence*, *1*, 00003. doi:10.1016/j.caeai.2020.100003

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers & Education: Artificial Intelligence*, *1*, 100001. doi:10.1016/j.caeai.2020.100001

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A Concept map-embedded educational computer game for improving students' learning performance in natural science courses, *Computers & Education*, *69*, 121-130. doi:10.1016/j.compedu.2013.07.008

Kalk, K., Luik, P., & Taimalu, M. (2019). The Characteristics of students, blog groups and blogging that predict reflection in blogs during teaching practice and induction year. *Teaching and Teacher Education*, *86*, 102900. doi:10.1016/j.tate.2019.102900

Krutka, D. G., Bergman, D. J., Flores, R., Mason, K., & Jack, A. R. (2014). Microblogging about teaching: Nurturing participatory cultures through collaborative online reflection with pre-service teachers. *Teaching and Teacher Education*, *40*, 83-93. doi:10.1016/j.tate.2014.02.002

Lau, S., Liem, A. D., & Nie, Y. (2008). Task-and self-related pathways to deep learning: The Mediating role of achievement goals, classroom attentiveness, and group participation. *British Journal of Educational Psychology*, *78*(4), 639-662.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

Lei, C., & Chan, C. K. (2018). Developing metadiscourse through reflective assessment in knowledge building environments. *Computers & Education*, *126*, 153-169. doi:10.1016/j.compedu.2018.07.006

Lim, J. Y., & Lim, K. Y. (2020). Co-regulation in collaborative learning: Grounded in achievement goal theory. *International Journal of Educational Research*, *103*, 101621. doi:10.1016/j.ijer.2020.101621

Liu, Q., Zhang, S., Wang, Q., & Chen, W. (2017). Mining online discussion data for understanding teachers reflective thinking. *IEEE Transactions on Learning Technologies*, *11*(2), 243-254.

Mercier, E. M. (2017). The Influence of achievement goals on collaborative interactions and knowledge convergence. *Learning and Instruction*, *50*, 31-43. doi:10.1016/j.learninstruc.2016.11.006

Meyer, K. A., & Xu, Y. J. (2007). A Bayesian analysis of the institutional and individual factors influencing faculty technology use. *The Internet and Higher Education*, *10*(3), 184-195.

Novakovich, J. (2016). Fostering critical thinking and reflection through blog-mediated peer feedback. *Journal of Computer Assisted Learning*, *32*(1), 16-30.

Ozkan, Y. (2019). Reflectivity of pre-service language teachers echoed through blogs. *Kasetsart Journal of Social Sciences*, *40*(1), 155-163.

Payne, S. C., Youngcourt, S. S., & Beaubien, J. M. (2007). A Meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology*, *92*(1), 128-150.

Qu, H., & Lee, H. (2011). Travelers' social identification and membership behaviors in online travel community. *Tourism Management*, *32*(6), 1262-1270.

Tajfel, H. (1978). The achievement of group identification. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the psychology of intergroup relations* (pp. 61-76). London, UK: Academic Press.

Thijs, J., & Fleischmann, F. (2015). Student–teacher relationships and achievement goal orientations: Examining student perceptions in an ethnically diverse sample. *Learning and Individual Differences*, *42*, 53-63. doi:10.1016/j.lindif.2015.08.014

Tikhomirova, T. S., & Kochetkov, N. V. (2018). Relationship between learning motivation and reflection in undergraduate students. *Psychological Science and Education*, *23*(6), 97-106.

Ullmann, T. D. (2019). Automated analysis of reflection in writing: Validating machine learning approaches. *International Journal of Artificial Intelligence in Education*, *29*(2), 217-257.

Urdan, T., & Kaplan, A. (2020). The Origins, evolution, and future directions of achievement goal theory. *Contemporary Educational Psychology*, *61*, 101862. doi:10.1016/j.cedpsych.2020.101862

Van den Broeck, A., Ferris, D. L., Chang, C. H., & Rosen, C. C. (2016). A Review of self-determination theory's basic psychological needs at work. *Journal of Management*, *42*(5), 1195-1229.

Van den Kieboom, L. A. (2013). Examining the mathematical knowledge for teaching involved in pre-service teachers' reflections. *Teaching and Teacher Education*, *35*, 146-156. doi:10.1016/j.tate.2013.06.009

Vygotsky, L. S. (1978). *Mind in society: The Development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Waheed, H., Hassan, S. U., Aljohani, N. R., Hardman, J., Alelyani, S., & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*, *104*, 106189. doi:10.1016/j.chb.2019.106189

Wang, X., Mei, X., Huang, Q., Han, Z., & Huang, C. (2021). Fine-grained learning performance prediction via adaptive sparse self-attention networks. *Information Sciences*, *545*, 223-240. doi:10.1016/j.ins.2020.08.017

Won, S., Wolters, C. A., & Mueller, S. A. (2018). Sense of belonging and self-regulated learning: Testing achievement goals as mediators. *The Journal of Experimental Education*, *86*(3), 402-418.

Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A Systematic review of journal publications from 2007 to 2017. *Computers & Education*, *140*, 103599. doi:10.1016/j.compedu.2019.103599

Yilmaz, R. (2016). Knowledge sharing behaviors in e-learning community: Exploring the role of academic self-efficacy and sense of community. *Computers in Human Behavior*, *63*, 373-382. doi:10.1016/j.chb.2016.05.055

Yilmaz, F. G. K., & Keser, H. (2016). The Impact of reflective thinking activities in e-learning: A Critical review of the empirical research. *Computers & Education*, *95*, 163-173. doi:10.1016/j.compedu.2016.01.006

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligenace Magazine*, *13*(3), 55-75.

Yukawa, J. (2006). Co-reflection in online learning: Collaborative critical thinking as narrative. *International Journal of Computer-Supported Collaborative Learning*, *1*(2), 203-228.

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A Review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, *31*(7), 1235-1270.

Zhou, Y., Huang, C., Hu, Q., Zhu, J., & Tang, Y. (2018). Personalized learning full-path recommendation model based on LSTM neural networks. *Information Sciences*, *444*, 135-152. doi:10.1016/j.ins.2018.02.053

Zou, D., Xie, H., Wang, F. L., & Kwan, R. (2020). Flipped learning with Wikipedia in higher education. *Studies in Higher Education*, *45*(5), 1026-1045.

Zumbrunn, S., McKim, C., Buhs, E., & Hawley, L. R. (2014). Support, belonging, motivation, and engagement in the college classroom: A mixed method study. *Instructional Science*, *42*(5), 661-684.