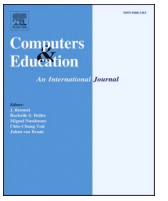
Computers & Education 191 (2022) 104644



Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu



Enhancing argumentative writing with automated feedback and social comparison nudging

Thiemo Wambsganss [a,*], Andreas Janson [b], Jan Marco Leimeister [b,c]

[a] EPFL, Machine Learning for Education Laboratory, Lausanne, Switzerland
[b] University of St. Gallen, Institute of Information Management, Switzerland
[c] University of Kassel, Research Center for IS Design, Germany

ARTICLE INFO

Keywords:
Argumentation
Automated feedback
Social comparison nudging
Natural language processing
Machine learning

ABSTRACT

The advantages offered by natural language processing (NLP) and machine learning enable students to receive automated feedback on their argumentation skills, independent of educator, time, and location. Although there is a growing amount of literature on formative argumentation feedback, empirical evidence on the effects of adaptive feedback mechanisms and novel NLP approaches to enhance argumentative writing remains scarce. To help fill this gap, the aim of the present study is to investigate whether automated feedback and social comparison nudging enable students to internalize and improve logical argumentation writing abilities in an undergraduate business course. We conducted a mixed-methods study to investigate the impact of argumentative writing on 71 students in a field experiment. Students in treatment group 1 completed their assignment while receiving automated feedback, whereas students in treatment group 2 completed the same assignment while receiving automated feedback with a social comparison nudge that indicated how other students performed on the same assignment. Students in the control group received generalized feedback based on rules of syntax. We found that participants who received automated argumentation feedback with a social comparison nudge wrote more convincing texts with higher-quality argumentation compared to the two benchmark groups ($p < 0.05$). The measured self-efficacy, perceived ease of use, and qualitative data provide valuable insights that help explain this effect. The results suggest that embedding automated feedback in combination with social comparison nudges enables students to increase their argumentative writing skills by triggering psychological processes. Receiving only automated feedback in the form of in-text argumentative highlighting without any further guidance appears not to significantly influence students' writing abilities when compared to syntactic feedback.

Thiemo Wambsganss, Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. Andreas Janson, Conceptualization, Methodology, Writing – review & editing, Supervision. Jan Marco Leimeister, Supervision.

* Corresponding author.
E-mail addresses: thiemo.wambsganss@epfl.ch (T. Wambsganss), andreas.janson@unisg.ch (A. Janson), leimeister@uni-kassel.de (J.M. Leimeister).

https://doi.org/10.1016/j.compedu.2022.104644
Received 13 March 2022; Received in revised form 14 August 2022; Accepted 19 September 2022
Available online 3 October 2022
0360-1315/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Argumentation is a fundamental component of everyday communication and thought (Kuhn, 1992; Toulmin, 2003). Convincing arguments are crucial for persuading an audience of a novel idea but also for negotiation, strategic decision-making, and civic dialogue (Eemeren et al., 1996; Kuhn, 1993; Scheuer, Loll, Pinkwart, & McLaren, 2010). However, research has demonstrated that humans are generally deficient in argumentation: they often fail to recognize the difference between merely expressing an opinion and making a fact-based claim (e.g., Byrne (1989); Marcus and Rips (1979); Tversky and Kahneman (1974)). Moreover, they often do not rebut others' arguments but ignore points of conflict and instead persist with their own arguments (Byrne, 1989). Students often struggle with developing argumentation skills due to a lack of individual and immediate feedback in their learning processes (Byrne, 1989; Eemeren et al., 1996; Hattie & Timperley, 2007; Marcus & Rips, 1979). Providing feedback on the individual argumentation skills of learners is time consuming and not scalable if conducted on a case-by-case basis by educators (Eemeren et al., 1996; Oecd, 2018; Wambsganss et al., 2020; Zhu, Liu, & Lee, 2020). Furthermore, novel distance learning scenarios such as massive open online courses (MOOCs) (Seaman, Allen, & Seaman, 2018) come with additional barriers related to individual feedback on a given learner's argumentation.

One possible solution to provide students with adaptive argumentation feedback in their learning processes is using recent developments in natural language processing (NLP). Methods from NLP and machine learning (ML) have been successfully used to predict argumentation quality in specific texts (Lawrence & Reed, 2019; Lippi & Torroni, 2015). In this context, research into argumentation mining is a crucial field for the development of learning support systems that identify arguments in unstructured texts (Lippi & Torroni, 2015). Argumentation mining can be used to score the quality of a text and provide tailored feedback regarding that text (Chernodub et al., 2019, pp. 195–200; Song, Heilman, Klebanov, & Deane, 2014; Stab & Gurevych, 2014a). For several decades, scientists and practitioners, especially in educational technology, have conceived and built tools to support the active teaching of argumentation by using input masks, representational guidance, or adaptive feedback to enhance students' learning of argumentation (e.g., De Groot et al. (2007); Osborne et al. (2016); Pinkwart, Ashley, Lynch, and Aleven (2009); Wambsganss et al. (2020); Afrin et al. (2021); Zhu et al. (2020); Lin, Fan, and Xie (2020); Yeh and She (2010); Valero Haro, Noroozi, Biemans, and Mulder (2019)).

While there is significant research into argumentation systems, only a few studies have investigated the specific effects of NLP-based automated writing feedback on areas like students' self-efficacy, revision behavior, or argumentation learning outcomes (Scheuer, 2015; Wambsganss et al., 2020). For instance, Latifi, Noroozi, Hatami, and Biemans (2021) developed an online platform called EduTech that enables students to receive instructional help when writing argumentative essays in a collaborative learning scenario. Their approach was not based on ML or NLP, but their findings did indicate that students in the scripted online peer feedback condition wrote essays with better argumentative writing quality than students in control groups. Afrin et al. (2021), meanwhile, empirically evaluated the impact of their system ArgRewrite, which provides students with adaptive argumentation self-monitoring through in-text highlighting. They found that detailed argumentation monitoring at the sentence level is more helpful than conditions that do not provide detailed feedback. However, their system was not based on argumentation mining but relied on Wizard of Oz-based human automation, in which a human would always provide students with argumentation feedback through the system. Zhu et al. (2020) investigate the role of adaptive NLP-based feedback on students' revision-taking behavior and scientific argumentation writing. They found that students' revisions were positively related to performance score increases and that adaptive feedback was more effective in assisting learning. However, the NLP-based scoring system employed for their adaptive feedback, c-rater-ML, is almost 10 years old and does not benefit from novel advances of argumentation mining since then (Zhu et al., 2020). In the domain of NLP and ML, a couple of systems have been demonstrated that make novel argumentation mining models available for users. However, none was designed for educational tasks or evaluated in an educational setting (e.g., Chernodub et al. (2019, pp. 195–200); Lippi and Torroni (2016); Lauscher, Glavaš, and Eckert (2018); Lawrence and Reed (2019); Stab and Gurevych (2017a)). Only a few studies have recently embedded argumentation mining into learning exercises to support students' acquisition of argumentation skills (Wambsganss et al., 2020; Wambsganss, Kueng, Söllner, & Leimeister, 2021). For example, Wambsganss et al. (2020) compiled a novel corpus of argumentation-annotated student peer reviews and embedded a trained argumentation mining model in the back end of a user-centered learning system called AL. They compared their method with a discuss-scripting approach (Fischer, Kollar, Stegmann, & Wecker, 2013) in a laboratory experiment with 54 students and found that students who received automated feedback wrote texts with a higher formal and perceived quality of argumentation following the argumentation measurement scheme of Weinberger and Fischer (2006). However, beyond initial proof-of-concept studies, the novel design and adoption of these automated argumentation learning systems lack further empirical evaluations, such as their possible impact on student learning and perception (Scheuer, 2015; Stab & Gurevych, 2017a). Field experiments that embed argumentation mining in theory-informed learning systems that are reusable and can be leveraged by large numbers of students are almost non-existent. Studies that provide further evidence for recently investigated positive effects of adaptive argumentation feedback based on argumentation mining remain rare (Wambsganss et al., 2020). In addition, detailed insights into the mechanisms of how to provide automated feedback on argumentation writing is also lacking, although these feedback processes are considered crucial to the successful deployment of digital learning systems (Söllner, Bitzer, Janson, & Leimeister, 2018).

In summary, the majority of existing empirical research in the area of automated feedback with argumentation mining and argumentation learning is rather explorative (e.g., Wizard of Oz, qualitative interviews with students and educators, and system demonstrations outside the educational domain), and only a few empirical studies rigorously measure the effectiveness of argumentation mining in argumentation learning environments exist (Wambsganss et al., 2020, 2021). This gap is particularly pronounced with respect to empirical research using field experiments in real-time learning environments and their effects on students' learning and perception, such as whether automated feedback with social comparison nudging helps students to train argumentation skills over

the course of an argumentative writing assignment conducted in the field.

Drawing on nudging theory as a guiding design principle (Thaler & Sunstein, 2009) to shed light on the mechanisms of providing more effective feedback, we aim to contribute to research by investigating, designing, and evaluating a novel student-centered learning system that provides students with automated feedback and a social comparison nudge irrespective of educator, time, and location (Bandura, 1991; Bjork, Dunlosky, & Kornell, 2013; Zimmerman & Schunk, 2001). To contribute to a better understanding of the effects of automated feedback on adaptive argumentation learning, this paper seeks to answer the following two research questions:

RQ1. Does automated argumentation feedback in combination with social comparison nudging help students develop argumentation writing skills in a persuasive writing assignment?

RQ2. How does automated argumentation feedback in combination with social comparison nudging influence students' learning processes?

To answer our research questions, we created an experimental design in which we complemented the measurement of students' argumentation writing skills (RQ1) with qualitative insights and learning perception measurements (e.g., self-efficacy) of the students (RQ2) in one field experiment. We investigate the impact of automated feedback and social comparison nudging on students' argumentation skills and their learning process during a persuasive writing task.

2. Theoretical background and related work

2.1. Educational technologies for argumentation learning

There is some evidence in prior research for the positive effects of formative feedback on training skills such as argumentation (Afrin et al., 2021; Hattie & Timperley, 2007; Jonassen & Kim, 2010; Van Eemeren, Van Eemeren, & Grootendorst, 2004; Wang, Arya, Novielli, Cheng, & Guo, 2020; Zhu et al., 2020). The output of feedback, according to Sadler (1989), is precise information relevant to the task or learning process that fills a gap between what is understood and what is sought to be comprehended. For argumentation skill learning, researchers have investigated adaptive support approaches (Scheuer et al., 2010) in which students receive pedagogical feedback about their actions, along with ideas and recommendations, to stimulate and lead their future writing activities. In most cases, a computer-based assessment is used to determine whether a given argument is syntactically and semantically correct (e.g., Pinkwart et al. (2009); Stab and Gurevych (2017b); Wambsganss et al. (2020); Afrin et al. (2021); Wang et al. (2020)). Automated feedback (which is also called individual, dynamic, adaptive, or personalized feedback) has been a particular area of interest for providing students with argumentation feedback on their skills (Afrin et al., 2021; Latifi et al., 2021; Noroozi, Kirschner, Biemans, & Mulder, 2018; Scheuer et al., 2010; Scheuer, Niebuhr, Dragon, McLaren, & Pinkwart, 2012; Valero, Noroozi, Biemans, & Mulder, 2022, pp. 123–145). However, the implementation of these educational ML-based systems is a complex endeavor that needs to be studied from an interdisciplinary perspective that draws on computer science, didactic design, and educational technology. Hence, as Scheuer (2015) and Stab and Gurevych (2017a) have noted, existing research lacks a rigorous empirical investigation of automated argumentation feedback approaches.

2.2. Argumentation feedback through digital nudging

To guide user interaction with feedback, researchers have successfully relied on the potential of digital nudging (Mirsch, Lehrer, & Jung, 2018; Thaler & Sunstein, 2009). The positive effects of nudging have previously been used in designing digital learning systems. For example, Motz, Mallon, and Quick (2021) found that push notifications about missing submissions nudged students into improved assignment adherence and course grades when compared to students who did not receive such nudges. Moreover, Eskreis-Winkler, Milkman, Gromet, and Duckworth (2019) reported that nudging students to give motivational advice to peers (e.g., how to stop procrastinating) leads to higher academic grades. However, to the best of our knowledge, the concept of nudging has not been investigated for argumentation skill learning with automated feedback, such as nudging users to write more persuasive texts based on social comparison (Cialdini & Goldstein, 2004; Mirsch, Lehrer, & Jung, 2017). The concept of social comparison (Cialdini & Goldstein, 2004), also called social nudging, is especially promising when combining learning feedback and goal setting (e.g., by providing a feedback message such as *"other users have used more arguments in the same task"*). For instance, nudges that include social comparisons with other learners have led to an increase in study time and better exam performance (O'Connell & Lang, 2018). In addition, it is important not to consider solely behavioral outcomes such as performance when researching AI-based argumentation learning (Scheuer et al., 2010; Valero Haro et al., 2019). Theoretically, the underlying mechanism for providing social comparison feedback is also rooted in social cognitive theory (Bandura, 1977) and is thus a prime candidate for positive behavior change in learning processes (Bandura, 1991), because the effective provision of feedback through digital nudging may positively influence learners' beliefs to conduct better arguments. Furthermore, as Al Shamsi, Al-Emran, and Shaalan (2022) have shown, system-related perceptions such as perceived ease of use related to AI-based learning systems are also crucial for learners when using the affordances provided by adaptive learning systems. Therefore, we aim to investigate whether and how automated argumentation feedback based on ML, combined with social comparison, will help students write more convincing texts. We believe that self-regulated learning theory supports our hypothesis that automated argumentation feedback increases students' self-efficacy and thus their learning of argumentation skills (Bandura, 1991; Zimmerman & Schunk, 2001).

2.3. Argumentation mining

To investigate our RQs, we designed a novel pedagogical scenario in which we provide students with automated feedback with a social comparison in a real-world writing exercise. To do so, we leverage argumentation mining, a relatively new research field in computational linguistics that focuses on the extraction and analysis of arguments from textual corpora. Over the last 15 years, scientists have published studies on argumentation mining in legal texts, online reviews, and debates (Cabrio & Villata, 2012; García-Villalba & Saint-Dizier, 2012; Lawrence & Reed, 2019; Mochales & Moens, 2011). Argumentation mining seeks to analyze the arguments of a given text based on a defined argumentation structure that often uses (Toulmin, 2003)). The identification of argumentation structures can be carried out on three levels: the first is to detect a sentence containing an argument to differentiate argumentative from non-argumentative text units (Florou, Konstantopoulos, Kukurikos, & Karampiperis, 2013); the second deals with classifying argument components into *claims* and *premises* (Mochales & Moens, 2011; Stab & Gurevych, 2014b); and the third is the identification of argumentative relations (Palau & Moens, 2009; Stab & Gurevych, 2014b). Researchers have shown increasing interest in intelligent writing assistance (Song et al., 2014; Stab & Gurevych, 2017a), since it enables adaptive argumentative learning with tailored feedback about arguments in texts (Lawrence & Reed, 2019; Lippi & Torroni, 2016). However, the complexity of using this technology in combination with social comparison in field interventions to provide students automated feedback has thus far been poorly assessed (Lawrence & Reed, 2019; Wambsganss et al., 2020). In our study approach, we focus on the first two subtasks to assess the argumentation level of students to provide automated feedback in combination with social comparison nudging (Mirsch et al., 2018).

3. Methods

3.1. Experimental context

We designed an assignment for higher education students that asked them to write a persuasive business pitch of 300 words (a translated extract from an exemplary student elevator pitch appears in appendix A). The study was conducted as a field experiment in a large-scale lecture course about information management at a university in Western Europe. In the lecture course in which we conducted that experiment, students developed and presented digital business models (Osterwalder & Pigneur, 2013). In one assignment, students were asked to submit a novel business idea based on a value proposition canvas; part of the assignment required writing a 300-word persuasive business pitch to convince potential investors of the value proposition of their idea. Taking part in the assignment was not mandatory for students to pass the class. However, by successfully completing the assignment, students could improve their final grade by 2.2%. The quality of the business pitch—that is, its persuasiveness—had no influence on the evaluation of the assignment and thus no influence on final grades. We asked students to write their business pitches on the Google Docs platform and provided them with different levels of feedback on their pitches' argumentation levels. However, we ensured that students also had the opportunity to complete the exercise using any other writing tool of their choice for data privacy reasons and due to our university's ethical standards. Students who chose to not use Google Docs were excluded from the sample.

We manipulated the way students received feedback on the argumentation level of their texts. Participants were randomly assigned to one of three groups: two treatment and one control (see Fig. 2). Students in the control group (CG) were provided with general argumentation feedback based on syntactic rules. Participants in treatment group 1 (TG1) received automated argumentation feedback including in-text highlighting based on a trained argumentation mining model. Participants in treatment group 2 (TG2) wrote the persuasive business pitch by receiving automated argumentation feedback in combination with a social comparison nudge based on the same trained argumentation mining model.

3.2. Participants

In total, 124 students from the course signed up for the assignment, of whom 71 successfully completed our experiment. After randomization, we counted 21 valid results in CG, 25 in TG1, and 25 in TG2 (see Table 1). CG participants had an average age of 24.17 (SD = 3.56, 9 males, 12 females), TG1 participants were (on average) 24.24 years old (SD = 3.83, 13 males, 12 females), and TG2 participants had an average age of 23.04 (SD = 2.32, 13 males, 12 females). The persuasive writing task took most participants between 30 and 45 min.

To ensure that the randomization resulted in randomized groups and were controlled for potential effects of interfering variables with our small sample size, we compared the differences in the means of personal innovativeness according to (Agarwal & Karahanna, 2000) and feedback-seeking of individuals (Ashford, 1986) included in the pre-test. For both constructs, we obtained p values larger than 0.05 between the TGs and the CG (for personal innovativeness, $p = 0.4697$; for feedback-seeking of individuals, $p = 0.5047$). This

Table 1
Overview of participants across three conditions.

Group	n	Age	Gender
TG2: automated feedback with social comparison nudge	25	mean = 23.04, SD = 2.32	13 males, 12 females
TG1: automated feedback	25	mean = 24.24, SD = 3.83	13 males, 12 females
CG: general syntactic feedback	21	mean = 24.17, SD = 3.56	9 males, 12 females

demonstrated that there was no significant difference in the mean values for these two constructs.

3.3. Learner interaction

We designed and built an adaptive learning tool that provides students with automated feedback on their argumentation skill level based on our argumentation mining model. In designing the tool, we followed the findings of recent user studies (Afrin et al., 2021; Wambsganss et al., 2020), a review of the extant literature related to the provision of feedback related to failure (e.g., Bjork et al. (2013); Metcalfe (2017); Lawrence and Reed (2019)), and self-regulated learning theory (Bandura, 1991; Zimmerman & Schunk, 2001). The learning tool is a Google Docs add-on that enables students to receive adaptive argumentation feedback whenever and wherever they want. Our goal is to provide learners with automated feedback that includes a social comparison nudge based on logical argumentation errors, irrespective of instructor, time, and location. Our system is illustrated in Fig. 1.

When starting the application, the students were introduced to the functionality of the tool and provided with a short overview of argumentation theory. Next, the learning tool displayed a simple feedback dashboard to the user. After users write their texts, they can select part of them and click on "Now analyze." Afterwards, they will receive automated argumentation feedback based on their argumentation skills. To prevent information overload, we tailored the function to only analyze selected paragraphs. The social comparison nudge counts the number of claims in the user's document and provides relative performance feedback (Damgaard & Nielsen, 2018) using historical data on the underlying knowledge base or task. For our assignment, we used the average of claims in the corpus of Wambsganss and Niklaus (2022). Prior students wrote an average of eight arguments (claims) in a 300-word pitch. If students had more claims than average, they received a positive feedback message (highlighted in an intuitive green), such as "Very nice. You displayed your idea with a sufficient number of arguments." However, if a student had fewer than eight claims, he or she received the following feedback message (highlighted in yellow as a hint): "Perhaps you could write some more arguments in your texts to improve the persuasiveness of your idea." We relied in our design on findings on digital nudging (Weinmann, Schneider, & Brocke, 2016) and that typically color codings for presenting social comparison nudges are perceived as rather subtle but often trigger effective behavioral changes through psychological processes, such as in our case, compared to the social norm.

3.4. Automated feedback algorithm

We develop a novel learning system enabling students to receive automated argumentation feedback with a social comparison. We relied on a corpus of 200 student-written business model pitches by Wambsganss and Niklaus (2022), trained an argumentation mining model based on the corpus annotations to predict the argumentation structures of students when writing a persuasive business pitch, and embedded the model in a theory-informed argumentation feedback system in Google Docs. During the argumentative writing exercise, students in TG1 used the feedback system and only received automated feedback, whereas students in TG2 used the same system with automated feedback that included a social comparison nudge.

We leverage the German business model pitch corpus of Wambsganss and Niklaus (2022), which consists of 200 pitches written by students extracted from a large-scale lecture scenario (an exemplary pitch with annotation from the corpus can be found in Fig. 3 in appendix A). The texts are annotated for their argumentative components (claims, premises) and for the relations between those components. Our objective was to embed a classification algorithm in the back end of an argumentative learning system to provide students with automated argumentation feedback in the writing process. The task is considered a sentence-based classification task in which each sentence can be either a *claim*, a *premise*, or *non-argumentative*. Therefore, we trained and tuned the hyperparameters of a long short-term memory (LSTM) model (Hochreiter & Schmidhuber, 1997) to classify the argumentative components of a given text, since several other authors have reached satisfactory results when using an LSTM for argumentation compound classification (e.g., Chernodub et al. (2019, pp. 195–200)). We tokenized the texts and transformed them into word embeddings (GloVe). The labels were one-hot encoded. The data set was split into training and test sets using an 80:20 stratified split. The LSTM architecture consisted of eight layers and a dropout rate of zero, with 60 units per layer. For component classification, we obtained an accuracy of 54.12%, precision of 55.90%, and a recall of 54.12% on the test data. We benchmarked our approach against a BERT model (Devlin, Chang, Lee, & Toutanova, 2019). For the BERT model, a learning rate of 1e-5, a batch size of 16, and training the model of over 25 epochs provided the best results. However, we obtained an accuracy of 47.50%, precision of 46.66%, and a recall of 47.50%, which were rather unsatisfactory.

Compared to a random classification (33.33% accuracy for a three-class prediction), our BERT model shows satisfactory results with an improvement of 21% (54% accuracy of our BERT model compared to 33% random). Moreover, we tested the accuracy of our automated feedback in a pre-study with a total of 37 students. The results from a qualitative and quantitative survey confirmed that our model appeared to be robust enough for students to receive satisfactory feedback on their argumentative texts.[1] Furthermore, we calculated two summary scores to provide students with an overview of the quality of their argumentation based on previously extracted argumentative structures, including *readability*, which is defined by how readable a text is based on the Flesch Reading Ease score (Flesch, 1943) and *argumentativeness*, which determines the proportion of argumentative and non-argumentative parts of a given text. These two rules of syntax were also the basis for calculating the general feedback for participants in the CG.

[1] The pre-study is published in Wambsganss, Janson, Käser, & Leimeister, 2022.

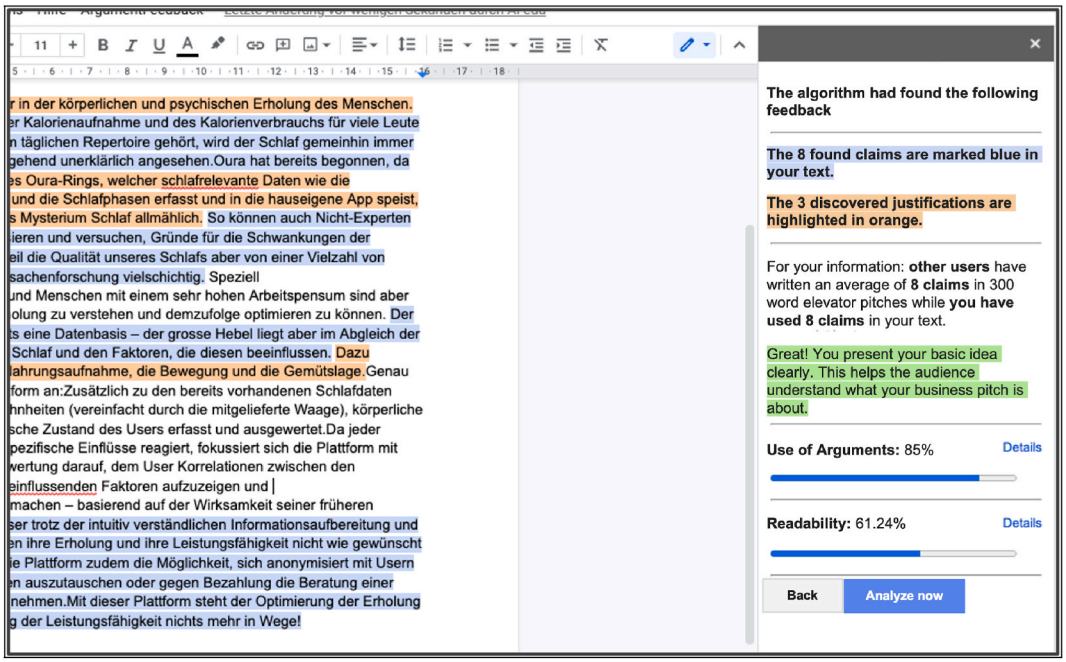


Fig. 1. Screenshot of the design of our automated argumentation feedback system featuring a social comparison nudge.

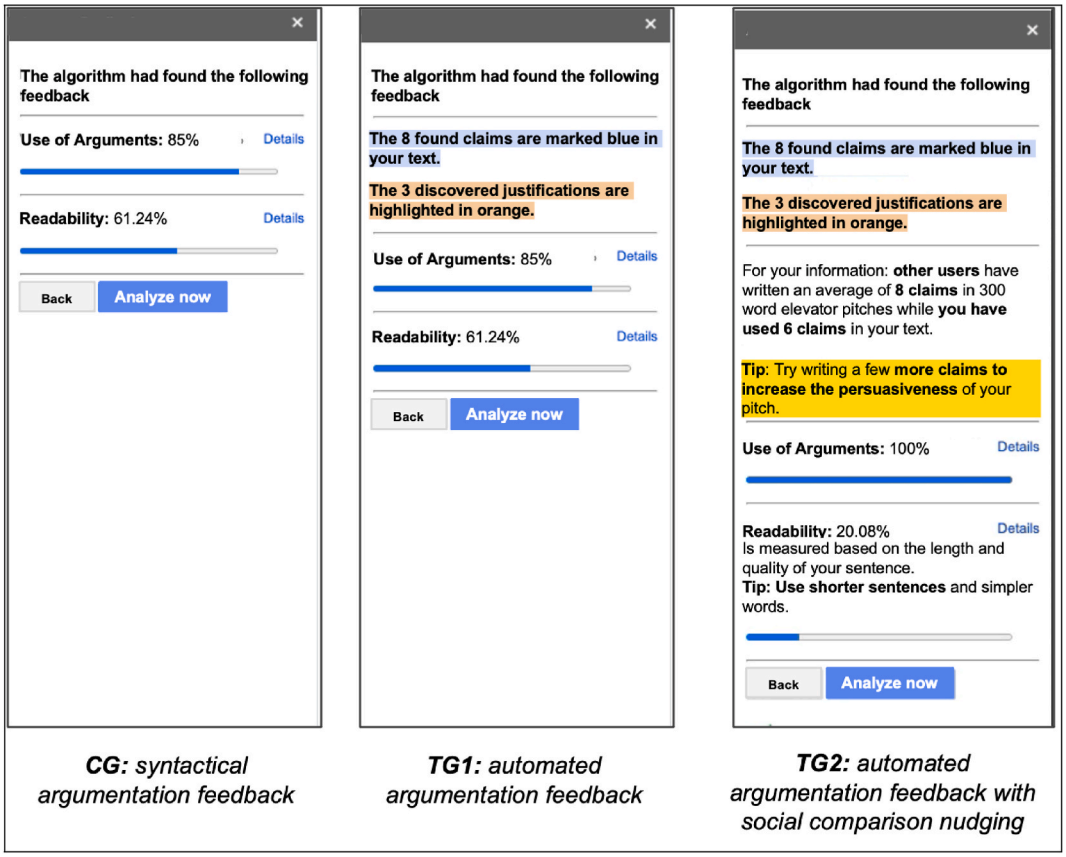


Fig. 2. Overview of the different groups.

3.5. Experimental procedure

The experiment consisted of three main parts: 1) a pre-test, 2) a persuasive writing exercise, and 3) a post-test. The pre- and post-phases were consistent for all participants. In the writing phase, TG1 participants received automated feedback, TG2 participants received automated feedback with a social comparison nudge to conduct the persuasive writing exercise, and CG participants conducted the same exercise with general argumentation feedback based on syntactic rules .[2]

1) **Pre-test:** The experiment started with a pre-survey of eight questions. We tested two different constructs to assess whether the randomization was successful. First, we asked four questions to test the participants' personal innovativeness in the domain of information technology, following (Agarwal & Karahanna, 2000). Second, we tested the construct of individuals' feedback-seeking, following (Ashford, 1986). Both constructs were measured on a 7-point Likert scale (1: totally disagree to 7: totally agree, with 4 a neutral statement). These items can be found in appendix B.
2) **Persuasive writing exercise:** In this stage of the experiments, we asked participants to write a persuasive business pitch of 300 words. Students were expected to create a succinct but persuasive pitch regarding the *"what, why, and how"* of their unique business idea. All students were instructed about (1) the basics of argumentative writing and argumentation theory based on the Toulmin model (Toulmin, 2003), (2) the characteristics of a sound persuasive pitch based on the value proposition canvas of Osterwalder and Pigneur (2013), and (3) the goal of the exercise ("write an argumentative pitch to convince potential investors"). Participants in TG1 received automated argumentation feedback including in-text highlighting argumentation when writing the pitch. Students in TG2 wrote the persuasive business pitch while receiving automated argumentation feedback in combination with a social comparison nudge, and CG participants conducted the same exercise while receiving general feedback based on syntactic rules.
3) **Post-test:** In the post-survey, we aimed to measure the perceived influence of our argumentation feedback on students' learning processes. To do so, we measured 1) the perceived ease of use of the argumentation feedback tool in students' learning processes based on Venkatesh and Bala (2008), 2) the self-efficacy of students for the task of argumentation skill learning based on three items, following Bandura (1991), and 3) included two items to capture the self-reported number of arguments in the students' final business pitches and their revision behavior by capturing how often they improved the argumentation level of their pitches based on the feedback received. Furthermore, we captured the qualitative impressions of students in their learning processes by asking three qualitative questions: "*What did you particularly like about the use of the argumentation tool?*", "*What could be improved?*", and "*Do you have any other ideas?*" Moreover, we incorporated one control question, included an item to serve as a manipulation check, and captured participant demographics. All quantitative constructs were measured on a 7-point Likert scale (1: totally disagree to 7: totally agree, with 4 a neutral statement).

3.6. Measurement of argumentation quality

Beyond measuring several subjective constructs, our main objective was to measure the argumentation quality of the written texts from the three groups. Therefore, the pitches were analyzed for their formal argumentation quality. We applied the annotation scheme for argumentative knowledge construction described by Weinberger and Fischer (2006). This annotation scheme has been applied in several studies and has demonstrated high objectivity, reliability, and validity (e.g., Stegmann, Wecker, Weinberger, and Fischer (2012); Wambsganss et al. (2020, 2021)). To measure the formal quality of the argumentation, the annotator had to distinguish between 1) *unsupported claims*, 2) supported *claims*, 3) *limited claims*, and 4) supported *and limited claims*. A more precise description of the scheme can be found in Weinberger and Fischer (2006). We thus trained two annotators based on the 16-page annotation guidelines in Wambsganss and Niklaus (2022) to assess the argumentation components of the pitches. The formal quality of the argumentation of each user was then defined by the number of arguments he or she wrote. Following Stegmann et al. (2012), only supported, *limited*, and supported *and limited claims* were counted as argumentation.

4. Results

We start by presenting the results related to our first research question: *Does automated argumentation feedback in combination with social comparison nudging help students develop argumentation writing skills in a persuasive writing assignment.* To investigate RQ1, we specifically tried to validate (1) whether automated argumentation feedback, compared to general feedback based on rules of syntax, helps students write better argumentative texts, and (2) whether adding social comparison nudging to automated feedback helps students write better argumentative texts.

4.1. Impact on argumentative writing

To answer RQ1, we compared the argumentation quality between the written texts of the CG and the TGs. We used an ANOVA analysis to check for significance between the variables. To ensure that the assumptions for the ANOVA model are met, we conducted statistical and visual tests for normality and homogeneity of variance (for more detail, see appendix C, Fig. 4). All assumptions were

[2] The exact items are all in appendix B in Table 4.

met for all variables. More specifically, for argumentation quality we conducted a Shapiro–Wilk test to test the normality assumption ($p = 0.12$) and a Levene's test to check for the homogeneity of variance assumption ($p = 0.09$). To measure the effects between the groups, we ran a Tukey post hoc test.

The average number of arguments in the persuasive writing exercise texts by participants receiving automated feedback with a social comparison nudge (TG2) was 4.87 (SD = 1.24). For texts by participants who received only automated feedback (TG1), we counted a mean of 3.9 arguments (SD = 1.69; see Table 2). People in the CG who received general feedback based on rules of syntax wrote texts with a mean of 3.64 arguments (SD = 1.20). The ANOVA analysis showed that there is a statistical difference in the formal argumentation quality of TG2 and TG1 students ($p = 0.04$ ($p < 0.05$)). We calculated a Cohen's d of 0.77, indicating a difference between the means of 0.7768 standard deviations. Since this effect is considered medium according to Cohen (1988), we assume that providing students automated feedback with social comparison nudging has a positive effect on their argumentation writing quality compared to general syntactic argumentation feedback. No significant effect was found between TG1 and CG as to formal argumentation quality ($p = 0.81$). This means that no general effect of automated argumentation feedback compared to general syntactic feedback was found. Hence, we have to state that providing automated feedback as opposed to syntactic feedback does not appear to automatically increase students' argumentation quality. However, we can confirm that automated feedback with a social comparison nudge does appear to increase the students' argumentation quality more than only automated feedback or syntactic feedback. The difference between TG1 and TG2 is statistically significant ($p < 0.01$). The Cohen's d of 0.5 shows that there is a medium effect between students receiving automated feedback with a social comparison nudge and students receiving only automated feedback.

4.2. Impact on students' learning processes

Since our findings for RQ1 lead us to assume that adding social comparison nudging to automated feedback helps students write better argumentative texts, we aimed to further investigate our RQ2 about *how automated feedback in combination with social comparison nudging influences students' learning processes* to obtain deeper insights into the observed effect and the mechanisms of argumentation feedback.

To do so, we combined 1) quantitative results of student perceptions through their argumentation learning process (such as perceived ease of use, self-efficacy, self-reported number of arguments, and argumentative revisions), and 2) the analyzed qualitative answers of students after they completed the argumentative writing exercises.[3] To evaluate the students' perceptions in their learning processes, we calculated the means for all groups for perceived ease of use. Surprisingly, we did not see any significant difference between the treatments (TG2: mean = 4.98, SD = 1.18; TG1: mean = 4.94, SD = 0.98; CG: mean = 5.33, SD = 0.71, $p > 0.05$). All values were normalized >0.7, indicating a positive ease of use of all treatments in students' learning processes (Venkatesh & Bala, 2008). In addition, we did not find a significant effect for self-efficacy between groups (TG2: mean = 4.77, SD = 0.71; TG1: mean = 4.98, SD = 0.98; CG: mean = 4.96, SD = 0.94, $p > 0.05$). In fact, the values for self-efficacy were all similarly very high when compared against the midpoints.

Comparing the self-reported argumentation and the reported argumentation revision behavior between the three groups (see Table 6 in appendix D), we found that students in TG2 reported 14.16 final arguments in their persuasive business pitches (SD = 5.63). An ANOVA with a Tukey post hoc test showed that this is significantly higher than what the students in the CG reported (mean of 8.0, SD = 4.11), with a p value of 0.003 ($p < 0.01$). Participants in TG1 reported a mean of 12.2 arguments (SD = 7.81). However, we did not find any significant difference between TG1 and TG2 or between TG1 and CG. We also compared the argumentative revisions that the participants reported between groups (see Table 5 in appendix D). However, we could not find any significance between groups in the distribution of the data. These results indicate that students who received automated feedback with a social comparison nudge were self-aware about the higher number of arguments in their texts. Nonetheless, they did not appear to change their argumentation more often based on the feedback. To further investigate this counterintuitive finding, we analyzed the qualitative responses from all students. We identified three main themes between the treatments. The topics involve 1) learning interaction, 2) feedback presentation, and 3) adaptivity and guidance. We briefly present the highlights and further elaborate on them in the discussion section. Moreover, we translated the responses from German and categorized the most representative response for TG2 in Table 3.

1) **Learning interaction:** Across all three groups, we see positive comments on the general interaction and the embedding of the learning intervention in the students' writing processes. Multiple students across the CG and both TGs highlighted the integration of the argumentation feedback tool into a natural working writing environment (like Google Docs) and positively cited the effect of the combination of argumentation theory and adaptive argumentation feedback on their learning processes. Interestingly, 81% of students from the CG (who received syntactic feedback based on two scores) also mentioned the positive effect of writing feedback and the benefit of percentage scores to motivate them to write more: for example, "I was fascinated by how a program could analyze my text and would like to use this tool for writing other texts, as I felt I could improve my own writing style and delete unnecessarily long phrases. Thus, I was able to reduce my text very well to the most important things and express my arguments to the point."
2) **Feedback presentation:** However, a difference in qualitative comments emerges when it comes to the presentation of feedback. 78% of students from TG1 and TG2 mentioned that they liked or were motivated by the in-text argumentation highlighting of both tools: for example, "I really liked the color marking and the handling regarding the quick analysis while writing." Several

[3] See Table 5 in appendix D for all quantitative results.

Table 2
Results on argumentation quality the groups on a 1–7 Likert scale (***p < 0.001, **p < 0.01, *p < 0.05).

Group	n	Mean	SD	p-value TukeyHSD Test TG1	p-value TukeyHSD Test (CG)
TG2: automated feedback with social comparison nudge	25	4.87	1.24	0.04* Cohen's d = 0.5 (medium effect)	0.01** Cohen's d = 0.77 (medium effect)
TG1: automated feedback	25	3.9	1.69	–	0.81
CG: general syntactic feedback	21	3.64	1.20	–	–

Table 3
Representative examples of qualitative user topics from TG2.

Topic	Exemplary User Response
Learning interaction	- **On general user interaction**, e.g., "ArgumentFeedback is simple and very easy to use. The integration into Google Docs specifically makes the argumentation feedback very handy." - **On interaction dynamics**, e.g., "In general, the tool should be more dynamic if possible. Going through an analysis every time is a bit cumbersome (tagging, etc.). Automatic updates of the feedback and metrics would improve the user experience a lot."
Feedback presentation	- **On social comparison nudges**, e.g., "The argumentative comparison with peers serves as an incentive to write at least as many or more arguments." -**On summary scores**, e.g., "Readability as a criterion is often neglected for argumentative texts. The feedback on this was very helpful for me to improve my pitch." - **On social comparison nudges**, e.g., "The argumentative comparison with peers serves as an incentive to write at least as many or more arguments." -**On social comparison nudges**, e.g., "The argumentative comparison with peers serves as an incentive to write at least as many or more arguments."
Adaptivity and guidance	-**On automated feedback**, e.g., "I was encouraged to write more arguments without stretching the text too much, as the ratio of the length of the text to the arguments is taken into account. Likewise, one is motivated to shorten the sentences. It is also good that tips are given for the categories in order to improve the evaluation." -**On accuracy**, e.g. "I'm not sure how well the tool recognized my arguments."

participants in TG1 (automated feedback without nudging) asked for additional guidance and feedback elements from the learning tool, such as "it would have been nice to have somewhat more specific feedback, such as sample sentences for improvement." Students from the CG even specified design elements: "Unfortunately, only two values were displayed to me: use of arguments and readability. Beyond that, however, little explanation of how these values were calculated (e.g., by highlighting in the text), so of course the learning factor was limited. Also, I did not get any statistics on the performance of other participants. That would have provided me with more guidance on what to improve," and "Markings on my text would have helped me know exactly which parts to improve."

3) **Adaptivity and guidance:** A difference was found in the comments on the adaptivity and guidance of the learning tool, especially regarding the feedback. Participants in TG2 (automated feedback with social nudging) cited social comparison as a motivating element to write more correct arguments: for example, "the argumentative comparison with peers serves as an incentive to write at least as many or more arguments." Although adaptivity of feedback was an improvement suggested across all groups, especially the CG, students also noted that the two scores could be more accurate.

5. Discussion of findings and theory implications

The objectives of the study were to investigate the effects of automated feedback on students' argumentation skills and how these interactions impact their learning processes. We found that the students who received automated feedback in combination with social comparison nudging wrote their texts with higher levels of argumentation skills. Against our expectations, students who only received automated feedback did not write better argumentative texts than students who received general feedback. This suggests that social comparison nudging in combination with automated feedback might be a valuable source for student learning. With the help of automated feedback and social comparison, students were able to receive formative feedback and individual support in a writing environment that they commonly use and were thus comfortable with. The quantitative and qualitative data acquired in our field experiment indicate a positive effect of automated argumentation feedback based on argumentation mining in combination with social comparison nudging. Interestingly, the perceived ease of use and self-efficacy were equally high in all three groups, showing the positive user acceptance of feedback learning tools in general. Hence, this study expands prior research around formative feedback and digital nudging that used mainly exploratory or non-automated approaches (e.g., Zhu et al. (2020); Afrin et al. (2021); Wambsganss et al. (2020); Lin et al. (2020)).

Although the quantitative data of our rather small sample do not allow for more detailed analysis, the qualitative data do suggest explanations for the effect of automated feedback in combination with social comparison. The results on perceived ease of use and self-efficacy indicate that automated feedback (whether based on syntactic rules or on ML and NLP algorithms) has a positive effect on student learning. This is also reflected in previous research that investigated the effect of automated feedback (often based on syntactic rules) in students' learning processes (e.g., Zhu et al. (2020); Afrin et al. (2021); Huang, Chang, Chen, Tseng, and Chien (2016)). These positive qualitative comments across all groups on the general concept and the embedding of an argumentation feedback tool in the

learning process provide a further indication of this sort. Students across all groups mentioned that they were motivated by the combination of an easy-to-use learning tool, which combined argumentation theory explanation, goal setting, and formative feedback. The high perceived ease of use and self-efficacy of all participants reflect and confirm the positive impact of design principles from prior literature (e.g., Wambsganss et al. (2020); Bandura (1991)).

The positive effect of social comparison on argumentative writing skills could be explained by nudging theory (Thaler & Sunstein, 2009). Social comparison appears to lead to a combination of learning feedback and goal setting (Cialdini & Goldstein, 2004). This is again reflected in the numerous qualitative comments we received from users about the peer comparison. Students appeared to be nudged in their argumentative writing behavior, as social comparisons with other learners led to better learning outcomes. This is also reflected in the literature about social conformity (Cialdini & Goldstein, 2004). Our study is in line with the results on social nudging for learning, such as O'Connell & Lang, 2018, who found that social comparison leads to an increase in study time and higher exam performance (O'Connell & Lang, 2018). The correct level of feedback on a student's skills, such as argumentation skills, appears to lead to high self-efficacy and thus better learning outcomes. Moreover, our results suggest that social comparison, with its intuitive and therefore subtle color coding according to digital nudging theory (Weinmann et al. (2016)), nudges students to write texts that are more argumentative when referring to the overall quality indicated through the quantification of qualified arguments. As detailed above, students perceived a generally positive experience in their learning processes across all feedback treatments. However, students who received automated feedback in combination with social comparison appeared to be nudged to write more formal arguments in their assignments. This is in line with prior literature on behavioral economics, which found that digital nudging often affects people's behavior in rather subtle ways (Thaler & Sunstein, 2009) but also underscores the importance of how to design social comparison nudging in education, although we want to highlight that our evidence did not directly suggest that the nudge is perceived as subtle. By providing real-world relative performance feedback from the existing corpus as a social comparison Damgaard and Nielsen (2018), students learned not only to focus on the quantity of arguments but also to construct high-quality arguments through the additional feedback of the system.

Our results appear to confirm the underlying mechanism rooted in social cognitive theory (Bandura, 1977) that automated feedback with social comparison leads to positive behavioral changes in learning processes (Bandura, 1991). For argumentation skill learning, this result is especially novel. Past research on adaptive argumentation learning tools based on computational methods, such as Afrin et al., 2021; Wambsganss et al., 2020; Wang et al., 2020, have focused largely on adaptive feedback like in-text highlighting (Afrin et al., 2021; Chernodub et al., 2019, pp. 195–200; Lippi & Torroni, 2016) or argumentation discourse monitoring (Wambsganss et al., 2020). Social comparisons for argumentative writing were previously largely neglected or even showed that users did not prefer to compare themselves with peers; see, for example, (Wambsganss & Rietsche, 2019. Moreover, traditional argumentation learning theory, such as the representational guidance theory in Suthers, 2003, focuses solely on argumentation monitoring; for example, by supporting argumentation learning through the provision of representations of argumentation structures with the objective to stimulate and improve individual reasoning, collaboration, and learning (Pinkwart et al., 2009; Suthers & Hundhausen, 2001). Our research provides novel insights into how to support students' argumentative writing based on automated feedback and digital nudging more efficiently.

Our study thus provides several contributions to research in educational technology. First, we contribute to computer tutoring research by providing empirical evidence on the effect of automated writing feedback on students' skill development. Second, we contribute in particular to the formative feedback literature stream by providing a nuanced perspective on different feedback mechanisms (syntactic feedback, automated feedback, and automated feedback with social comparison nudging) on students' learning processes and learning outcomes. To the best of our knowledge, empirical evidence for automated argumentation feedback in combination with a digital nudge is still limited in this area. Third, we contribute to technology-mediated learning by showing that interactions with a writing feedback tool in a common writing environment have the potential to impact students' learning processes, resulting in increased levels of skill development. The qualitative results indicate that this new technology might be able to offer automated feedback for self-regulated learning in a more natural and adaptive way. Especially in distance-learning scenarios such as MOOCs and the mass lectures that are common at public universities, automated writing feedback in common writing environments such as Google Docs could be a scalable, beneficial addition to current learning environments. Finally, we provide a contribution to design-based research on educational tools. As Wang and Hannafin (2005) highlighted, with our paper we also contribute principles for implementing writing support systems for other researchers to investigate the effect of these systems on students learning outcomes and/or interactions.

We believe our findings also offer several practical implications for researchers and educators to design general educational scenarios. First, we provide an archetypal use case of how a metacognitive skill (in this instance, argumentation) can be trained at scale in a typical pedagogical scenario. The power of adaptive writing feedback based on NLP has also been shown for other metacognitive skills like empathy (e.g., Wambsganss, Soellner, Koedinger, and Leimeister (2022)) and problem solving (e.g., Winkler, Söllner, and Leimeister (2021)). In this regard, this study shows educators how to build and integrate an adaptive writing tool based on digital nudging in an existing learning environment. Future researchers can follow a similar methodology to embed existing NLP technologies (e.g., emotion analysis) or other widely available argumentation mining approaches (for available corpora and a review, see Lawrence and Reed (2019)) to train students, as with scenarios about communication skills.

6. Limitations and future work

However, our work does have certain limitations. First, the accuracy of our argumentation mining algorithm leaves room for improvement. We have two basic options to improve the underlying model's performance: 1) by enriching the corpus used with more

annotated business pitches or 2) by improving the performance of the model itself. Our objective is to improve the model by annotating the written business pitches that we collected in our experiment and adding them to the corpus available from prior research (Wambsganss & Niklaus, 2022). Second, our study was limited in sample size. Further empirical evaluations are necessary to replicate the results. Third, although, we received a satisfying Cronbach's alpha (>0.7) for all perceptual constructs, the measured constructs of self-efficacy and ease of use come with the natural limitations of self-reported measurements (Taber, 2018). Fourth, we have only provided insights into the short-term effects of automated feedback. In our experiment, we showed that automated feedback with social comparison has a short-term influence on a student's argumentation skills. For future work, we suggest measuring the long-term learning effects on students' skills of different nudging interventions. With regards to measuring, we also note that future research could look into how social comparison nudging could improve individual argument quality and quality of overall argumentation with additional measures instead. By this means, we also acknowledge as a limitation of the present study that the overall argumentation quality is limited by counting qualified arguments. Future research could provide an even deeper quality perspective and shed light on how nudges contribute to multiple argumentation quality dimensions such as specificity, relevance, or clarity (Ke, Inamdar, Lin, and Ng (2019); Wambsganss and Niklaus (2022)). Fifth, we want to highlight the ethical limitations of our study. Regarding the implementation of automated feedback in a learning system, we do not want to replace human educators by any means, as we believe that skilled teachers will always be able to provide better individual writing feedback than any algorithm. However, we hope that our system can help human educators focus more time and attention on detailed questions and difficult cases. We also recognize several data privacy concerns with integrating our tool in a cloud-based writing editor, since personal data, such as students' argumentation skill levels, might be exposed to large technology companies. Hence, we call for future discussions on how to overcome the trade-off between making novel user-centered learning applications widely accessible and easy to use, perhaps by integrating them into known cloud environments without exposing learner data to third parties.

Finally, we call for future research regarding the effects of adaptive writing feedback on different user groups. Noroozi et al. (2022), for example, have found that female students provided better justifications for problems identified in peer review, more constructive reviews, and higher-quality peer review overall than males. Although we did not find any gender differences for any variable in our experiment, it might be interesting to explore the effect of adaptive writing feedback in combination with social nudging on different demographic variables to control for bias and fairness in our algorithmic training and tool design. In addition, in our research we only considered the effect of adaptive feedback on argumentation skills. Valero et al. (2022, pp. 123–145) found a significant relationship between students' argumentation behaviors and domain-specific knowledge acquisition. Hence, we call for future research to examine the effects of adaptive argumentation feedback in general and in combination with social nudging on domain-specific knowledge creation. Moreover, while automated feedback and social comparison nudging may have a positive effect on the majority of students, it may also be demotivating for some minority groups, e.g., for lower performing students. Although, we did not see any negative effects in our sample size, future research should shed additional light on unintended consequences of automated feedback in combination with social comparison.

All in all, we seek to offer insights into learner-centered design to further improve educational applications based on advances in computational methods. As NLP and ML progress further, we hope our work will attract researchers to design more intelligent forms of educational systems for other learning scenarios or metacognition skills and thus contribute to the OECD Learning Framework 2030 and its goal of a metacognition-skill-based education.

7. Conclusion

We have answered our two research questions. First, we investigated whether automated feedback with social comparison improved students' argumentative writing skills. In a field experiment in a higher education setting, we observed that the provision of automated argumentation feedback with a social comparison nudge to students had a positive effect on their argumentative writing skills. Second, we investigated how the use of automated feedback with social comparison nudging influenced students' learning processes. Students reported a high ease of use and self-efficacy when using the automated feedback approach. Our results suggest that social comparison nudges students to write more argumentative texts by triggering basic psychological processes, i.e., comparing and adapting to the social norm. Our research provides empirical evidence for the effectiveness of ML-based automated feedback in combination with digital nudging and offers fresh insights into how NLP- and ML-based feedback can support students' learning processes in a writing exercise.

Acknowledgements

We express our gratitude to the students of the University of Kassel who took part in this study. We would also like to thank Daniel Fuchs for his support in the development of the intervention and support in the pre-studies.

Funding

We thank the Swiss National Science Foundation for funding parts of this research (100013_192718 \& 200207). The second author acknowledges funding from the Basic Research Fund (GFF) of the University of St. Gallen.

Ethical approval

In this study, all procedures involving human participants were in accordance with the ethical standards of institutional and/or national research committees.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A

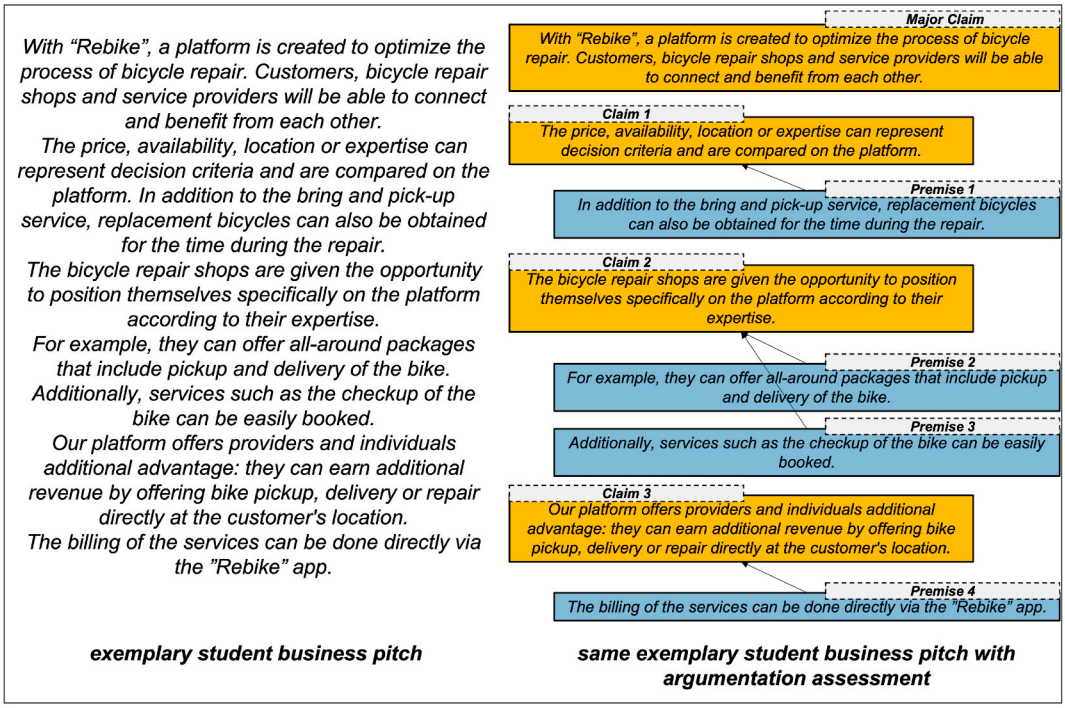


Fig. 3. Overview of an exemplary student business pitch translated based on the example in Wambsganss and Niklaus (2022). Left side, original pitch: right side, annotated pitch.

Appendix B

Table 4
Overview of items measured in pre- and post-tests.

Test	Variable	Items	Scale
Pre-test	personal innovativeness in the domain of information technology (Agarwal & Karahanna, 2000)	*"I like to experiment with new information technologies." "If I heard about a new information technology, I would look for ways to experiment with it." "In general, I am hesitant to try out new information technologies." "Among my peers, I am usually the first to try out new information technologies"*	1–7 (7: highest)
Pre-test	feedback-seeking individuals (Ashford, 1986)	*"It is important for me to receive feedback on my performance." "I find feedback on my performance useful." "I would like to get more feedback on what behaviors would help me better conduct a task." "It is important to me to receive feedback on my progress potential for skill learning."*	1–7 (7: highest)

(*continued on next page*)

Table 4 (*continued*)

Test	Variable	Items	Scale
Persuasive writing exercise	assignment	Write an argumentative pitch (approx. 300 words), which shows the value proposition of your elaborated business model. The elevator pitch should contain as many convincing arguments as possible.	
Post-test	manipulation check	With which functions did you just get feedback from ArgumentFeedback?	Condition 1, 2, or 3
Post-test	perceived ease of use (Venkatesh & Bala, 2008)	*"It would be easy for me to become adept at using the reasoning tool." "I find the reasoning tool easy to interact with." "Learning how to use the reasoning tool would be easy for me."*	1–7 (7: highest)
Post-test	attention check	*Please select "strongly agree."*	1–7 (7: highest)
Post-test	self-efficacy (Bandura, 1977)	*"In comparison to other users, I will write a good argumentative text." "I am sure that I could write a very good argumentative text." "I think I now know quite a bit about argumentative writing."*	1–7 (7: highest)
Post-test	attention check	*Please select "strongly agree."*	1–7
Post-test	self-reported number of arguments	*"How many arguments did you use in your text?"*	open
Post-test	argumentative revisions	*"How many times did you revise the argumentation in your pitch?"*	open
Post-test	qualitative impression	*"What did you particularly like about the use of the argumentation tool?" "What could be improved?" "Do you have any other ideas?"*	open
Post-test	pre-experience	*"Have you used a writing support program before (e.g., Grammarly)?"*	Yes/No
Post-test	demographics	1. Age 2. Gender	open

Appendix C

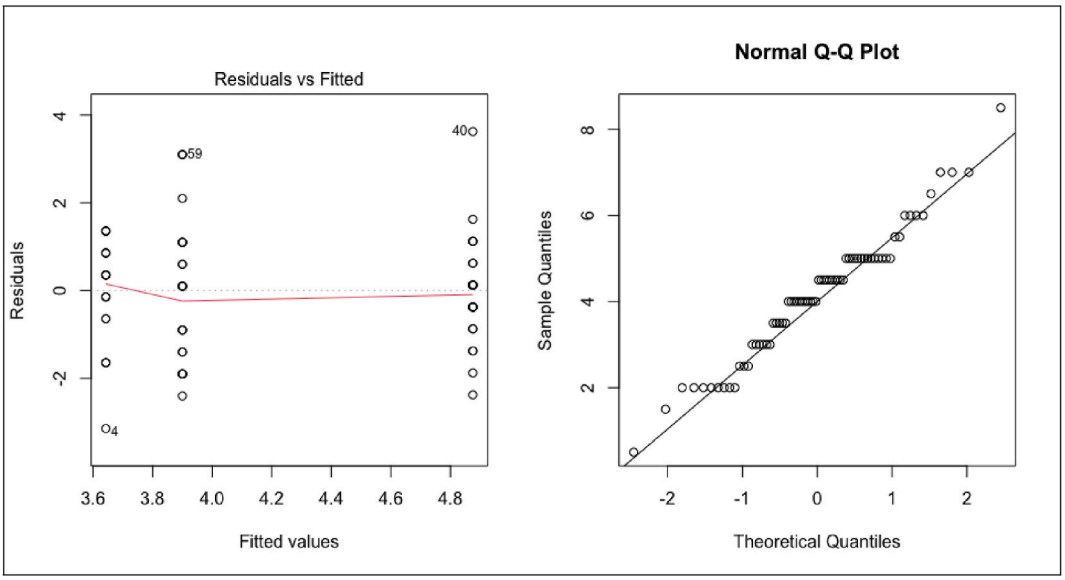


Fig. 4. Graphical check for assumptions for ANOVA for argumentation writing quality.

Appendix D

Table 5
Results of items measured in pre- and post-tests (α = Cronbach's alpha).

Variable	Group	Mean	
personal innovativeness, $\alpha = 0.76$	TG2: automated feedback with social comparison nudge	4.3 (SD = 0.73)	
	TG1: automated feedback	4.29 (SD = 0.56)	
	CG: general syntactic feedback	4.46 (SD = 0.57)	no significant difference ($p > 0.05$)
feedback-seeking individual, $\alpha = 0.83$	TG2: automated feedback with social comparison nudge	5.6 (SD = 0.82)	
	TG1: automated feedback	6 (SD = 0.68)	
	CG: general syntactic feedback	5.84 (SD = 0.8)	no significant difference ($p > 0.05$)
perceived ease of use, $\alpha = 0.79$	TG2: automated feedback with social comparison nudge	4.98 (SD = 1.18)	
	TG1: automated feedback	4.94 (SD = 0.98)	
	CG: general syntactic feedback	5.33 (SD = 1.02)	no significant difference ($p > 0.05$)
self-efficacy, $\alpha = 0.73$	TG2: automated feedback with social comparison nudge	4.77 (SD = 0.71)	
	TG1: automated feedback	4.98 (SD = 0.98)	
	CG: general syntactic feedback	4.96 (SD = 0.94)	no significant difference ($p > 0.05$)
argumentative revisions	TG2: automated feedback with social comparison nudge	4.08 (SD = 3.18)	
	TG1: automated feedback	3.28 (SD = 3.44)	
	CG: general syntactic feedback	3.33 (SD = 2.33)	no significant difference ($p > 0.05$)
pre-experience		no significant difference ($p > 0.05$)	
Age	TG2: automated feedback with social comparison nudge	23.04 (SD = 2.32)	
	TG1: automated feedback	24.24(SD = 3.83)	
	CG: general syntactic feedback	24.17 (SD = 3.56)	no significant difference ($p > 0.05$)
Gender	TG2: automated feedback with social comparison nudge	13 males, 12 females	
	TG1: automated feedback	13 males, 12 females	
	CG: general syntactic feedback	9 males, 12 females	no significant difference ($p > 0.05$)
Mother tongue			no significant difference ($p > 0.05$)

Appendix E

Table 6
Results of self-reported argumentation between the groups (***$p < 0.001$, **$p < 0.01$ *$p < 0.05$).

Group	Mean	SD	p value TukeyHSD Test TG1	p value TukeyHSD Test (CG)
TG2: automated feedback with social comparison nudge	14.16	5.63	0.5	0.003**
TG1: automated feedback	12.2	7.81	–	0.06
CG: general syntactic feedback	8.0	4.11	–	–

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2022.104644.

References

Afrin, T., Kashefi, O., Olshefski, C., Litman, D., Hwa, R., & Godley, A. (2021). Effective interfaces for student-driven revision sessions for argumentative writing. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–13). New York, NY, USA: Acm. URL: https://dl.acm.org/doi/10.1145/3411764.3445683. doi:10.1145/3411764.3445683.

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly, 24*, 665. URL: https://www.jstor.org/stable/3250951?origin=crossref. doi:10.2307/3250951.

Al Shamsi, J. H., Al-Emran, M., & Shaalan, K. (2022). *Understanding key drivers affecting students' use of artificial intelligence-based voice assistants URL* https://doi.org/10.1007/s10639-022-10947-3. doi:10.1007/s10639-022-10947-3.

Ashford, S. (1986). Feedback-seeking in individual adaptation: A resource perspective. *Academy of Management Journal, 29*, 465–487. URL: http://amj.aom.org/cgi/doi/10.2307/256219. doi:10.2307/256219.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215. https://doi.org/10.1037/0033-295x.84.2.191. URL:/record/1977-25733-001.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*, 248–287. https://doi.org/10.1016/0749-5978(91)90022-l

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Byrne, R. M. J. (1989). Everyday reasoning with conditional sequences. *The Quarterly Journal of Experimental Psychology Section A, 41*, 141–166. URL: http://journals.sagepub.com/doi/10.1080/14640748908402357. doi:10.1080/14640748908402357.

Cabrio, E., & Villata, S. (2012). natural language arguments: A combined approach. In *20th European conference on artificial intelligence (ECAI 2012)*. Montpellier, France.

Chernodub, A., Oliynyk, O., Heidenreich, P., Bondarenko, A., Hagen, M., Biemann, C., et al. (2019). *Targer: Neural argument mining at your fingertips*, 10.18653/v1/pp.19–3031.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge. https://doi.org/10.4324/9780203771587

Damgaard, M. T., & Nielsen, H. (2018). Nudging in education. *Economics of Education Review, 64*, 313–342. URL: https://EconPapers.repec.org/RePEc:eee:ecoedu:v:64:y:2018:i:c. :p:313-342.

De Groot, R., Drachman, R., Hever, R., Schwarz, B. B., Hoppe, U., Harrer, A., et al. (2007). Computer supported moderation of e-discussions: The argunaut approach. In *Proceedings of the 8th iternational conference on computer supported collaborative learning* (pp. 168–170). International Society of the Learning Sciences.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the north American chapter of the association for computational linguistics: Human language technologies, ume 1* pp. 4171–4186). Minneapolis, Minnesota: (Long and Short Papers), Association for Computational Linguistics. URL: https://aclanthology.org/N19-1423. doi:10.18653/v1/N19-1423.

Eemeren, F. H.v., Grootendorst, R., et al.Grootendorst, R., Johnson, R. H., Plantin, C., & Willard, C. A. (1996). *Fundamentals of argumentation theory*. Routledge. URL: https://www.taylorfrancis.com/books/9780203811306. doi:10.4324/9780203811306.

Eskreis-Winkler, L., Milkman, K. L., Gromet, D. M., & Duckworth, A. L. (2019). A large-scale field experiment shows giving advice improves academic outcomes for the advisor. In *Proceedings of the national academy of sciences of the United States of America* (Vol. 116, pp. 14808–14810). URL: www.pnas.org/cgi/doi/10.1073/pnas.1908779116. doi:10.1073/pnas.1908779116.

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist, 48*, 56–66. URL: http://www.ncbi.nlm.nih.gov/pubmed/23378679 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3557614. doi:10.1080/00461520.2012.748005.

Flesch, R. (1943). Marks of readable style; a study in adult education. *Teachers College Contributions to Education, 897*.

Florou, E., Konstantopoulos, S., Kukurikos, A., & Karampiperis, P. (2013). Argument extraction for supporting public policy formulation. In *Proceedings of the 7th workshop on language technology for cultural heritage, social sciences* (pp. 49–54). Acl: and Humanities.

García-Villalba, M. P., & Saint-Dizier, P. (2012). A framework to extract arguments in opinion texts. *International Journal of Cognitive Informatics and Natural Intelligence, 6*, 62–87. URL: http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/jcini.2012070104. doi:10.4018/jcini.2012070104.

Hattie, J., & Timperley, H. (2007). The power of feedback. URL: http://www.columbia.edu/mvp19/ETF/Feedback.pdf.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*, 1735–1780. URL: http://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf, doi:10.1162/neco.1997.9.8.1735.

Huang, C. J., Chang, S. C., Chen, H. M., Tseng, J. H., & Chien, S. Y. (2016). A group intelligence-based asynchronous argumentation learning-assistance platform. *Interactive Learning Environments, 24*, 1408–1427, 10.1080/10494820.2015.1016533.

Jonassen, D. H., & Kim, B. (2010). Arguing to learn and learning to argue: Design justifications and guidelines. *Educational Technology Research & Development, 58*, 439–457. https://doi.org/10.1007/s11423-009-9143-8

Ke, Z., Inamdar, H., Lin, H., & Ng, V. (2019). Give me more feedback II: Annotating thesis strength and related attributes in student essays. In *Proceedings of the 57th annual meeting of the association for computational linguistics* (pp. 3994–4004). Florence, Italy: Association for Computational Linguistics. URL: https://aclanthology.org/P19-1390. doi:10.18653/v1/P19-1390.

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review, 62*, 155–179. http://hepgjournals.org/doi/10.17763/haer.62.2.9r424r0113t670l1, doi:10.17763/haer.62.2.9r424r0113t670l1.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77*, 319–337. https://doi.org/10.1002/sce.3730770306. http://doi.wiley.com/10.1002/sce.3730770306.

Latifi, S., Noroozi, O., Hatami, J., & Biemans, H. J. (2021). *How does online peer feedback improve argumentative essay writing and learning?, 58* pp. 195–206). https://doi.org/10.1080/14703297.2019.1687005. doi:10.1080/14703297.2019.1687005. publisher: Routledge _eprint: https://doi.org/10.1080/14703297.2019.1687005.

Lauscher, A., Glavaš, G., & Eckert, K. (2018). ArguminSci: A tool for analyzing argumentation and rhetorical aspects in scientific writing. In *Proceedings of the 5th workshop on argument mining* (pp. 22–28). Brussels, Belgium: Association for Computational Linguistics. URL: https://aclanthology.org/W18-5203. doi:10.18653/v1/W18-5203.

Lawrence, J., & Reed, C. (2019). Argument mining: A survey. *Computational Linguistics, 45*, 765–818, 10.1162/COLIa00364.

Lin, Y. R., Fan, B., & Xie, K. (2020). The influence of a web-based learning environment on low achievers' science argumentation. *Computers & Education, 151*, Article 103860. URL: https://www.sciencedirect.com/science/article/pii/S0360131520300609. doi:https://doi.org/10.1016/j.compedu.2020.103860.

Lippi, M., & Torroni, P. (2015). Argumentation mining: State of the art and emerging trends. *IJCAI International Joint Conference on Artificial Intelligence 2015-Janua,*, 4207–4211, 10.1145/2850417.

Lippi, M., & Torroni, P. (2016). Margot: A web server for argumentation mining. *Expert Systems with Applications, 65*, 292–303. https://doi.org/10.1016/j.eswa.2016.08.050. doi:10.1016/j.eswa.2016.08.050.

Marcus, S. L., & Rips, L. J. (1979). Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior, 18*, 199–223. https://doi.org/10.1016/s0022-5371(79)90127-0. URL:/record/1980-27128-001.

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology, 68*, 465–489. https://doi.org/10.1146/annurev-psych-010416-044022

Mirsch, T., Lehrer, C., & Jung, R. (2017). Digital nudging: Altering user behavior in digital environments. In *Proceedings der 13*. Internationalen Tagung Wirtschaftsinformatik (WI 2017).

Mirsch, T., Lehrer, C., & Jung, R. (2018). Making digital nudging applicable: The digital nudge design method. In *International conference on information systems 2018*. ICIS 2018, Association for Information Systems. AIS Electronic Library (AISeL). URL: https://research.cbs.dk/en/publications/making-digital-nudging-applicable-the-digital-nudge-design-method.

Mochales, R., & Moens, M. F. (2011). *Argumentation mining. Artificial intelligence and law 19, 1–22*. URL: http://link.springer.com/10.1007/s10506-010-9104-x. doi:10.1007/s10506-010-9104-x.

Motz, B., Mallon, M. G., & Quick, J. D. (2021). Automated educative nudges to reduce missed assignments in college. *IEEE Transactions on Learning Technologies*. https://doi.org/10.1109/tlt.2021.3064613

Noroozi, O., Banihashem, S. K., Taghizadeh Kerman, N., Parvaneh Akhteh Khaneh, M., Babayi, M., Ashrafi, H., et al. (2022). *Gender differences in students' argumentative essay writing, peer review performance and uptake in online learning environments 0, 1–15*. https://doi.org/10.1080/10494820.2022.2034887. doi:10.1080/10494820.2022.2034887. publisher: Routledge _eprint: https://doi.org/10.1080/10494820.2022.2034887.

Noroozi, O., Kirschner, P. A., Biemans, H. J., & Mulder, M. (2018). Promoting argumentation competence: Extending from first- to second-order scaffolding through adaptive fading. *Educational Psychology Review, 30*, 153–176. https://link.springer.com/article/10.1007/s10648-017-9400-z. doi:10.1007/s10648-017-9400-z.

O'Connell, S. D., & Lang, G. (2018). Can personalized nudges improve learning in hybrid classes? Experimental evidence from an introductory undergraduate course. *Journal of Research on Technology in Education, 50*, 105–119. https://www.tandfonline.com/doi/abs/10.1080/15391523.2017.1408438. doi:10.1080/15391523.2017.1408438.

Oecd. (2018). *The future of education and skills - education 2030*. . doi:2018-06-15.

Osborne, J. F., Henderson, J. B., MacPherson, A., Szu, E., Wild, A., & Yao, S. Y. (2016). The development and validation of a learning progression for argumentation in science. *Journal of Research in Science Teaching, 53*, 821–846, 10.1002/tea.21316.

Osterwalder, A., & Pigneur, Y. (2013). Designing business models and similar strategic objects: The contribution of is. *Journal of the Association for Information Systems, 14*, 237–244, 10.17705/1jais.00333.

Palau, R. M., & Moens, M. F. (2009). Argumentation mining: The detection, classification and structure of arguments in text. *Proceedings of the 12th international conference on artificial intelligence and law*, 98–107. https://doi.org/10.1145/1568234.1568246

Pinkwart, N., Ashley, K., Lynch, C., & Aleven, V. (2009). Evaluating an intelligent tutoring system for making legal arguments with hypotheticals. *International Journal of Artificial Intelligence in Education, 19*, 401–424. http://iaiedsoc.org/pub/1302/file/19%5F4%5F05%5FPinkwart.pdf.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144. http://link.springer.com/10.1007/BF00117714. doi:10.1007/bf00117714.

Scheuer, O. (2015). *Towards adaptive argumentation learning systems*. Ph.D. thesis. URL: https://www.researchgate.net/publication/298087259.

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning, 5*, 43–102. https://doi.org/10.1007/s11412-009-9080-x

Scheuer, O., Niebuhr, S., Dragon, T., McLaren, B. M., & Pinkwart, N. (2012). Adaptive support for graphical argumentation – the LASAD approach. *IEEE Learning Technology Newsletter, 14*, 8–11. https://www.dfki.de/en/web/research/projects-and-publications/publications-overview/publication/8612/.

Seaman, J. E., Allen, I. E., & Seaman, J. (2018). Higher education reports - babson survey research group. Technical Report. URL: http://www.onlinelearningsurvey.com/highered.html https://www.onlinelearningsurvey.com/highered.html.

Söllner, M., Bitzer, P., Janson, A., & Leimeister, J. M. (2018). Process is king: Evaluating the performance of technology-mediated learning in vocational software training. *Journal of Information Technology, 33*, 233–253. http://journals.sagepub.com/doi/10.1057/s41265-017-0046-6. doi:10.1057/s41265-017-0046-6.

Song, Y., Heilman, M., Klebanov, B. B., & Deane, P. (2014). Applying argumentation schemes for essay scoring. In *Proceedings of the first workshop on argumentation mining* (pp. 69–78).

Stab, C., & Gurevych, I. (2014a). Annotating argument components and relations in persuasive essays. In *Proceedings of COLING 2014, the 25th international conference on computational linguistics: Technical papers* (pp. 1501–1510). http://www.ukp.tu-darmstadt.de.

Stab, C., & Gurevych, I. (2014b). Annotating argument components and relations in persuasive essays. In *Proceedings of the 25th international conference on computational linguistics* (pp. 1501–1510). COLING 2014).

Stab, C., & Gurevych, I. (2017a). Parsing argumentation structures in persuasive essays. *Computational Linguistics, 43*, 619–659. http://www.mitpressjournals.org/doi/abs/10.1162/COLI%5Fa%5F00295. doi:10.1162/COLI_a_00295.

Stab, C., & Gurevych, I. (2017b). Recognizing insufficiently supported arguments in argumentative essays. Technical Report. URL: www.ukp.tu-darmstadt.de.

Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. *Instructional Science, 40*, 297–323. https://doi.org/10.1007/s11251-011-9174-5

Suthers, D. D. (2003). Representational guidance for collaborative inquiry. In *Arguing to learn* (pp. 27–46). Springer. https://www.researchgate.net/publication/200772687. doi:10.1007/978-94-017-0781-7_2.

Suthers, D. D., & Hundhausen, C. D. (2001). European perspectives on computer-supported collaborative learning. Technical Report. URL: http://lilt.ics.hawaii.edu/papers/2001/Suthers-Hundhausen-Euro-CSCL-2001.pdf.

Taber, K.S.. (2018). *The use of cronbach's alpha when developing and reporting research instruments in science education* (Vol. 48, pp. 1273–1296). https://doi.org/10.1007/s11165-016-9602-2. doi:10.1007/s11165-016-9602-2.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press. https://doi.org/10.1016/s1477-3880(15)30073-6

Toulmin, S. E. (2003). *The uses of argument* (Updated edition). Cambridge University Press. doi:10.1017/cbo9780511840005.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.

Valero Haro, A., Noroozi, O., Biemans, H., & Mulder, M. (2019). *First- and second-order scaffolding of argumentation competence and domain-specific knowledge acquisition: A systematic review, 28* pp. 329–345). https://doi.org/10.1080/1475939X.2019.1612772. doi:10.1080/1475939X.2019.1612772. publisher: Routledge _eprint: https://doi.org/10.1080/1475939X.2019.1612772.

Valero, H. A., Noroozi, O., Biemans, H., & Mulder, M. (2022). *Argumentation competence: Students' argumentation knowledge, behavior and attitude and their relationships with domain-specific knowledge acquisition* (Vol. 35). URL: https://www.tandfonline.com/doi/citedby/10.1080/10720537.2020.1734995, doi:10.1080/10720537.2020.1734995. publisher: Routledge.

Van Eemeren, F., Van Eemeren, F. H., & Grootendorst, R. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge University Press.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*, 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x. http://doi.wiley.com/10.1111/j.1540-5915.2008.00192.x.

Wambsganss, T., Kueng, T., Söllner, M., & Leimeister, J. M. (2021). ArgueTutor: An adaptive dialog-based learning system for argumentation skills. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–14). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3411764.3445781.

Wambsganss, T., & Niklaus, C. (2022). *Modeling persuasive discourse to adaptively support students' argumentative writing. 60th annual meeting of the association for computational linguistics, ACL 2022 - proceedings of the conference*.

Wambsganss, T., Niklaus, C., Cetto, M., Söllner, M., Leimeister, J. M., & Handschuh, S. (2020). AL: An adaptive learning support system for argumentation skills. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–14). New York, NY, USA: Association for Computing Machinery.

Wambsganss, T., & Rietsche, R. (2019). Towards designing an adaptive argumentation learning tool. In *40th international conference on information systems* (pp. 1–12). Munich, Germany: ICIS 2019.

Wambsganss, T., Soellner, M., Koedinger, K. R., & Leimeister, J. M. (2022a). Adaptive empathy learning support in peer review scenarios. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3491102.3517740. doi:10.1145/3491102.3517740.

Wambsganss, Thiemo, Janson, Andreas, Käser, Tanja, & Leimeister, Jan Marco (2022b). Improving students argumentation learning with adaptive self-evaluation nudging. In *Proceedings of the ACM on Human-Computer Interaction* (vol. 6)CSCW2. https://doi.org/10.1145/3555633. ISSN 2573–0142.

Wang, W., Arya, D., Novielli, N., Cheng, J., & Guo, J. L. (2020). ArguLens: Anatomy of community opinions on usability issues using argumentation models. In *ACM CHI conference on human factors in computing systems*. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3313831.3376218.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research & Development, 53*, 5–23. http://www.jstor.org/stable/30221206.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education, 46*, 71–95. https://doi.org/10.1016/j.compedu.2005.04.003

Weinmann, M., Schneider, C., & Brocke, J.v. (2016). *Digital nudging. Business and information systems engineering* (Vol. 58, pp. 433–436). https://link.springer.com/article/10.1007/s12599-016-0453-1. doi:10.1007/s12599-016-0453-1.

Winkler, R., Söllner, M., & Leimeister, J. M. (2021). Enhancing problem-solving skills with smart personal assistant technology. *Computers & Education, 165*. https://doi.org/10.1016/j.compedu.2021.104148

Yeh, K. H., & She, H. C. (2010). On-line synchronous scientific argumentation learning: Nurturing students' argumentation ability and conceptual change in science context. *Computers & Education, 55*, 586–602. URL: https://www.sciencedirect.com/science/article/pii/S0360131510000631. doi:https://doi.org/10.1016/j.compedu.2010.02.020.

Zhu, M., Liu, O. L., & Lee, H. S. (2020). The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. *Computers & Education, 143*, Article 103668. URL: https://www.sciencedirect.com/science/article/pii/S0360131519302210. doi:https://doi.org/10.1016/j.compedu.2019.103668.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*. Routledge.