International Journal of Artificial Intelligence in Education (2021) 31:638–679
https://doi.org/10.1007/s40593-020-00204-4

ARTICLE

Collaborating with Mature English Language Learners to Combine Peer and Automated Feedback: a User-Centered Approach to Designing Writing Support

Amna Liaqat[1] · Cosmin Munteanu[2] · Carrie Demmans Epp[3]

Published online: 20 July 2020

Abstract

300,000 immigrants move to Canada each year in search of better economic opportunities, and many have limited English language skills. Improving written literacy of newcomers can enhance education, employment, or social integration opportunities. However, frequent, timely, and personalized feedback is not always possible for immigrants. Online writing support tools can scaffold writing development by providing this feedback, but existing systems provide inadequate support when instructors are inaccessible. In this paper, we show how feedback system design can leverage peer and automated feedback to support mature English Language Learners' (ELL) needs and practices. We identify strong associations between epistemic beliefs and learning strategies, highlighting the importance of tasks that activate productive epistemic beliefs. We find learners accurately assessed high-level issues in a peer's writing and are accepting of automated feedback, demonstrating that a platform combining peer-review and machine feedback could promote meaningful discussions. We present the results of our mixed-methods investigation that integrates three sources of information: analysis of learners' psychometric constructs, writing samples to identify error patterns, and participatory design group sessions incorporating human-centred design methods. We synthesize our results into four guidelines derived from seven findings resulting from the investigation of a system that scaffolds writing development for mature immigrant ELLs in the absence of formal instructional support. First, we find that ELLs require a platform to collaboratively iterate through the writing process. Next, we suggest how peer feedback can be enhanced through automated support. We then demonstrate how rubric design can guide both linear and holistic peer-review. Finally, we illustrate why open learner models and learning dashboards should contextualize real world progress.

✉ Amna Liaqat
aliaqat@cs.toronto.edu

[1] University of Toronto, Ontario, Canada

[2] University of Toronto Mississauga, Ontario, Canada

[3] University of Alberta, Alberta, Canada

Keywords Writing support · English language learners · Migrants · Human-centred design · Participatory design

Mature English Language Learners

Of the approximately 272,000 immigrants to Canada in 2015, 62% intended to find work (Immigration, Refugees and Citizenship Canada 2016). For 91% of all immigrants, English was not their native language, with 23% reporting no English language ability. In total, over 20% of the current population was born outside Canada. Recognizing the importance of language in adapting to a new country, the Government of Canada offers free English Language classes, called Language Instruction for Newcomers to Canada (LINC), to new immigrants (Immigration, Refugees and Citizenship Canada 2013). These classes may last several weeks to a few months, and they are offered full or part time. While Canada is an officially bilingual country and LINC offers support for both official languages, this work was conducted in one of the English-speaking regions so we focus on English language development.

Effective written communication is an essential but often underdeveloped skill for success in school and the workplace. Resource constraints in LINC, as well as traditional classrooms, limit the amount of one-on-one, individualized attention teachers can provide to students. However, improving written literacy, especially for adult migrants who speak English as an additional language, can lead to better employment, education, or social opportunities (Derwing et al. 2012). This relationship has been thoroughly investigated in countries, such as Canada, which see a large annual influx of migrants. It is also reasonable to expect that improving written literacy for mature immigrant English Language Learners (ELLs) is beneficial in other socio-economic settings.[1]

Research on improving written literacy or Learning to Write (LTW) has generally focused on understanding the motivations and challenges faced by post-secondary learners writing within an academic context. It has done so by analyzing writing rhetoric, errors, and feedback in order to design analytics and models that guide learners towards making meaningful revisions (Gibson et al. 2017; Perin and Lauterbach 2018; Shum et al. 2016; Ullmann 2017) or by designing help-seeking systems for learners ("Peerceptiv—Data Driven Peer Assessment n.d.). Both ELLs and older adults outside of the university setting have received comparatively little attention. Relative to young, native English speakers, much less is understood about what motivates mature immigrant ELLs (those between the ages of 30 to 60) to improve their written literacy, the challenges they face in learning to write, and how the design of writing tools can provide formative feedback that scaffolds these learners towards making meaningful revisions to their writing.

This study builds an initial understanding within this space by identifying the motivations, beliefs, challenges, and needs of mature immigrant ELLs who are improving their written English. We build on existing peer-learning research by employing user-centered design methods to engage with mature ELLs. We explore

[1] Throughout this paper, when we refer to our participants, we imply mature immigrant ELLs, although for brevity we may use only the term ELLs.

how to personalize the learning experience of mature immigrant ELLs, as personalization is crucial to providing timely help and guiding sense-making (Brooks et al. 2014; Julita Vassileva et al. 2016). A two-phase in-depth mixed-methods investigation with 15 adult ELLs was conducted to identify the goals of these learners, the processes they use to achieve them, and the design requirements for a support tool. From this, we suggest four design guidelines to help developers of language learning tools that are intended to support this underserved population. First, a platform that allows mature ELLs to collaboratively improve their writing skills would benefit both reviewer and writer (DG1). Second, a platform that incorporates automated feedback could scaffold learners through the technical aspects of writing tasks (DG2). Third, rubrics for peer-review should be deconstructed into manageable tasks that allow both linear and holistic review (DG3). Fourth, open learner models and learning dashboards may motivate consistent practice through real-world contextualization of progress (DG4). We expand prior research on peer support and open learner models (Bull and Kay 2010; Greer et al. 2001; Zapata-Rivera and Greer 2002) to a population whose needs have been underserved. This expansion integrates peer review with the types of feedback that are typically seen in open learner models to support student reflection over their abilities and help them identify which aspects of their writing they need to work on if they are going to move beyond their current skill set.

Understanding Learners and Assessing Writing

Providing just-in-time, formative feedback to learners is crucial for writing skills development. Understanding the writing challenges faced by mature ELLs as well as the individual goals, beliefs, motivations, and strategies that influence their behaviour is required to inform the design of writing support tools for this population. Though there is a large body of work on writing feedback for English language learners, the findings are often conflicting and much debated (Shao 2015; Waller and Papi 2017). One explanation of these differing findings is that writing feedback research generally focuses on overall trends instead of individual learner profiles, creating a gap in our understanding of individual learner differences. (Ferris 2010; Shao 2015). Unique learner needs, strengths, and weaknesses must be understood in order to provide automated, individualized support similar to the personalized feedback that instructors provide intuitively to their students (Ferris 2010). This need for understanding and adapting to individual differences is especially pronounced among underserved populations such as mature, migrant ELLs, where there is a need to understand how this group differs from other learners. In this literature review, we first provide an overview of existing writing support tools and the type of feedback or support they provide. The limitations of these, as they relate to adult ELLs, are discussed. Next, we discuss the various psychometric constructs that can affect a learner's approach to the writing process.

Writing Assessment and Feedback Delivery

Many writing support systems aim to assess writing quality and provide actionable feedback to learners. These systems generally have two major components. One, they

analyze learner writing. This analysis may include an instructor or it may rely on automated tools that generate analytics or create a model of the learner's knowledge and skills. Two, these systems use this assessment to deliver feedback to the writer, hoping that the feedback is comprehendible and meaningful. This feedback provisioning is consistent with much of what is seen in learning analytics dashboards and open learner models (Bodily et al., 2018; Bull 2020) - hereafter jointly referred to as open learner models (OLM).

Automated Assessment and Feedback

Automated components of writing support systems may analyze writing according to a set of rules or a model that was generated using data-driven approaches. These approaches automate assessment by mining the system's repository of submitted writing to identify patterns that can be used to tailor feedback for future submissions. These patterns can then be used to provide formative feedback to learners as long as models are interpretable, (H. Zhang et al. 2019) as is the goal with open learner models (Bull 2020; Bull and Kay 2016; Kay 1999). With Academic Writing Analytics (AWA), sequences of labels in high and low quality writing can be compared to discover patterns and generate actionable feedback that provides assessment on rhetorical quality to students as they write (Gibson et al. 2017). Like AWA, the potential for actionable feedback with Elouazizi et al.'s (2017) system lies in benchmarking a learner's essay against some standard and prompting learners to revise their writing to mimic the features of this standard. Similarly, Writing Pal (W-Pal) uses natural language processing (NLP) tools to provide automated formative feedback that is based on the analytics output by those NLP tools (Roscoe and McNamara 2013).

Human Assessment and Feedback

In contrast, Peerceptiv is a cloud-based peer-review system with a conceptually different approach to developing writing support tools. Peerceptiv does not provide automated feedback, instead all feedback is provided by peer-reviewers (Peerceptiv—Data Driven Peer Assessment n.d.). In a similar fashion, the PHelpS and iHelp systems provided peer support (Julita Vassileva et al. 2016). However, they did so in different contexts, with PHelpS being used in a work environment and iHelp used to support the understanding of computer science concepts and writing or debugging of code in post-secondary settings.

Writing Support Systems for Mature ELLs

Most computer-based writing support tools have been designed to assess writing in academic contexts with native speakers. Comparatively little attention has been given to ELLs, with most research involving this group having been done in academic contexts (Mathews-Aydinli 2008) rather than the everyday settings in which adult ELLs usually find themselves, such as searching for employment (Immigration, Refugees and Citizenship Canada 2016). When considering the appropriateness of existing systems for mature immigrant ELLs, a major limitation of using these tools is that they were designed for younger students in academic contexts.

W-Pal was developed with high-school students, while the other tools discussed were designed for post-secondary students who are highly proficient users of English. This may not appear to be an issue at first, but the features of these systems can make them challenging to use beyond the context for which they were built. One associated challenge is that low proficiency writers have greater diversity in the types of errors they produce, which makes designing targeted interventions more challenging (Perin and Lauterbach 2018). Also, the strategies, rhetoric, and linguistics of ELL writing differ from that of native speakers (Crossley and McNamara 2009; Silva 1993). ELLs tend to plan less before they write and produce structurally simpler writing (Silva 1993). ELLs also make use of different rhetorical devices. A common rhetorical error made by ELLs, in general, is when they attempt to translate a phrase word-by-word from their native language into English, resulting in a syntactically incorrect translation (Connor 1996).

ELLs may also struggle to understand the typically technical language used to provide feedback based on automated assessment (such as in AWA). This is even more problematic when there is no instructor to provide context. Many of these tools require an instructor to facilitate the software. However, many mature ELLs may not be taking classes or have access to an instructor to facilitate this writing development process. This lack of timely access to an instructor is one of the motivations behind the iHelp system (Bull et al. 2001) and our work here. These requirements and features make such systems impracticable for immigrant ELLs who wish to develop their skills outside of formal learning environments.

Group and individual differences must factor into the design of language learning tools for marginalized populations. For example, ELLs practicing their spoken literacy seek opportunities to learn in context. This practice can be supported by a search function that allows users to access context-specific vocabulary (Demmans Epp 2018). Marginalized populations may also have certain beliefs and perceptions that influence their learning process. Low-literacy adults and ELLs prefer to practice in private to avoid what they perceive as negative attention. For these users, a tool that can be used without overtly signalling the user as a language learner is particularly valuable (Demmans Epp 2018; Munteanu et al. 2014). When working with such groups, this understanding of both group and individual differences in user practices, beliefs, and needs can be uncovered through user-centered design approaches (Demmans Epp 2017).

Understanding Learners: Goals, Beliefs, Emotions, and Strategies

Understanding a learners' achievement goal orientations, epistemic beliefs and emotions, and cognitive strategies is essential for designing interventions to scaffold learners towards their goals. Interventions can encourage learners to adapt their strategies to meet their goals. However, such persuasive strategies are mediated by learner's beliefs and attitudes, and must be tailored to the learner to promote desirable learning outcomes (Orji et al. 2019). According to Krashen's affective filter hypothesis, language learning is influenced by affective factors such as motivation and anxiety (Krashen 1983). Studying the affective, motivational, and cognitive aspects of feedback at the same time brings together multiple dimensions of the learning process, which some have argued should be jointly studied (Shute 2008). In addition to individual variance, cultural factors (e.g., learners in an individualist context are influenced more

by reward and less by social comparison) mediate persuasive strategies and their effectiveness in influencing behaviour and attitudes (Orji et al. 2019). These potential differences between learners suggest a need for personalization within and across tasks or platforms, which would require user models that bridge contexts and user needs, as was the case with the Massive User Modelling System (MUMS) (Brooks et al. 2004).

Consistent with this general learner need, the design of a writing tool for adult, migrant ELLs should be informed by an understanding of their goals, beliefs, emotions, and learning strategies, as these constructs guide learning behaviour (Krashen 1983; Silva 1993). Muis (2007) proposed an integrated model combining goal orientations, epistemic beliefs, and learning strategies. A*chievement goal orientations*, which are cognitive conceptualizations of what an individual aims to achieve, affect the strategies learners' adopt and how they evaluate their performance (Pintrich et al. 2003). Individuals with a mastery goal orientation aim to learn the content of the task, while those with performance goals tend to focus on their ability relative to peers or some other external assessment (Ames 1992; Pintrich 2000; Pintrich et al. 2003).

A learner's achievement goal orientations are associated with their beliefs about the nature of knowing (Schommer 1990). These e*pistemic beliefs* are associated with the cognitive strategies that learners use during a task. Muis' (2007) empirically tested model showed that learners' epistemic beliefs influence the *learning strategies* they used, including how they defined task requirements, their method and goals for completing the task, the cognitive strategies used to reach those goals, and the internal standards set to evaluate the task after completion (K. R. Muis and Franco 2009). Learners also experience *epistemic emotions* when engaging in a task. These emotions arise from the cognitive aspects of a task, such as curiosity, boredom, frustration, and confusion and can greatly influence learning and performance (D'Mello and Graesser 2012; Pekrun et al. 2017). Negative feelings like frustration and boredom may lead to task abandonment (D'Mello and Graesser 2012) or ELLs disengaging from the types of activities that are known to support language learning (Demmans Epp 2016).

Epistemic beliefs and emotions are not fixed. Rather, they are domain-specific and can change depending on the task (K. R. Muis et al. 2006) and ELL context (Demmans Epp 2016). Domain-specific beliefs, as they relate to writing and particularly, for older adult ELLs have not been well explored. A 2009 survey of over 40 landmark studies in writing assessment only reported two studies involving adults (Cumming 2009), and neither of these studies investigated mature ELLs learning in non-academic contexts (Cumming 2009). Adults ELLs face challenges that have not been explored in previous writing support tool research, such as different motivations and goals for learning to write, and these exclusive features of ELL writing have not been explicitly represented in currently existing software. Moreover, contextual and individual differences make the results of such studies non-transferable to adult ELLs learning in non-academic contexts (Mathews-Aydinli 2008).

To enable the development of appropriate adaptive tools for this population, we must examine the links between the domain-specific, epistemic writing beliefs and emotions of adult ELLs. Doing so will allow us to make sense of their motivations and learning strategies to better design feedback systems for this population. With a more concrete understanding of the various underlying psychometrics at play during the writing process, recommendations can be made for feedback systems that promote on-task behavior, engagement, and learning.

Study Overview and Objectives

We conducted an investigation with 15 mature ELLs who are recent immigrants to Canada to better understand the needs of mature immigrant ELLs learning to write while collecting data to inform the development of analytic tools. This sample size is typical of design-based human-computer interaction (HCI) research (Caine 2016). The methodology was modelled after other studies that work with non-traditional populations where sample sizes may be smaller and analysis of qualitative data is required (Paulus 1999; Yang et al. 2006).

This design-based, mixed-methods study was performed in two phases. The first phase involved understanding the characteristics, needs, and motivations of mature ELLs learning to write using psychometric scales and interviews. During this phase, we also analyzed writing samples, which consisted of essays and peer-feedback on those essays. This analysis aimed to identify common patterns and errors faced by mature immigrant learners. To supplement the results from the first phase, the second phase explored the requirements of a tool that allows immigrant learners to independently engage in writing exercises and receive peer feedback. We collaborated with these mature ELLs using Participatory Design (PD), an approach from human-centered design methodologies that involves users throughout the design process (Vines et al. 2013).

Qualitative research is useful in revealing a deeper understanding about users, which cannot always be captured with larger scale, quantitative studies. Such an understanding is needed to inform the design of technologies when quantitative methods may not be as revealing (Baecker et al. 1995; Card et al. 1983; Moran 1996). Qualitative HCI is complementary to other quantitative methods, and it can be used at all stages of design (Baecker et al. 1995; Gould et al. 1991).

Given that our focus is gaining an understanding about a specific user group, the above methodology was considered appropriate, as opposed to formulating an a priori hypothesis and verifying it through statistical testing. Such an approach is suitable for understanding underrepresented users in a learning context, especially when the larger goal is designing technology that better addresses their needs (Munteanu et al. 2012). In this work, we expand on existing research on peer-learning by engaging with mature ELLs learning to write through user-centered design methods. We focus in this paper on understanding individual learner differences to inform the design of adaptive support for writing tools (Ferris 2010; Waller and Papi 2017). The result of this process is a set of guidelines for advising developers in the design of language learning tools for this underserved population.

Phase One: Understanding Learners and Assessing Writing

Associations between epistemic beliefs and other learning factors have not been well studied for learners in general, and even less so for mature immigrant ELLs. To address this lack of understanding we set two objectives for this phase. The first objective was to observe the role of various psychometric constructs in participants' writing process. Participants completed questionnaires that measured their achievement goals as well as their motivation and epistemic beliefs. The second objective aimed to discover common themes across learners' writing profiles. Participants provided writing samples and

engaged in peer-review. They were also presented with feedback they believed to be from several sources to capture their epistemic emotions throughout the peer-review process.

Participants

We used purposive sampling for participant recruitment (Etikan 2016) because we wanted to understand the particular and unique learning needs of a specific population. For this, we identified learning environments that aim to address the needs of our target population. As such, participants in this study were recruited from the Language Instruction for Newcomers to Canada (LINC) program. LINC is a government funded program offering free English language classes to recent immigrants. All students are assessed before being placed in a LINC class. Students are placed in courses based on this assessment, with courses dedicated to each of the proficiency levels defined by the assessment instrument.

LINC uses an assessment that is tied to the Canadian Language Benchmarks (CLB) standard, which is a scale describing language proficiency. CLB is divided into three stages. Individuals in stage one can use language in basic and predictable contexts. In stage two, individuals can participate in a variety of contexts and independently engage in routine and familiar situations. In the third stage, individuals can use language in sophisticated and high-stakes contexts (Hajer and Kaskens 2012). At the time of the study, all participants were enrolled in a stage two LINC class so they had similar English speaking, writing, and reading proficiency.

A researcher visited LINC classes in a large, predominantly English-speaking metropolitan area in Canada to invite students to participate in the study. Participants were informed during recruitment that they would receive $50 CAD and reimbursement for travel expenses to study sites. The researcher was not affiliated with LINC. Both the researcher and the program coordinators facilitating recruitment made it clear that participation was completely voluntary and not a component of their LINC class. The study was approved by the university's research ethics board.

While sample sizes vary for similar qualitative studies across the HCI field (Caine 2016), we stopped enrollment once we reached 15 participants. This number was deemed satisfactory because the second part of our study involved participatory design which tends to engage users in longer sessions and produce large amounts of data (Vines et al. 2012). Of the 15 ELLs who participated in this study, 11 were female. The uneven gender split in this study is representative of the split in LINC classes, where 72% of students are female (Government of Canada 2011). Participant ages ranged from 31 to 59 years ($M = 40.1$, $SD = 9.2$). Two participants immigrated under protected status (e.g., refugees), while the rest were either sponsored or skilled immigrants, which is the predominant migration class in Canada. All participants were either unemployed or employed part-time. For almost all participants (93%), getting a job or returning to school was the main motivator for taking LINC classes. All participants had completed at least some post-secondary training before moving to Canada. Four had either a master's degree or a PhD. This sample reflects the Canadian skilled immigration program, which favours highly-educated migrants who are selected using a competitive point-based system. The skilled worker program requires that immigrants demonstrate a certain level of competency in either official language (English or French) but not

both. This is reflected in our participant pool, with English being a second language for 12 participants and a third language for three of them. The prior education and language abilities of participants indicates that they are likely to succeed and are highly capable learners: high literacy in one or more languages allows one to draw on resources from other languages which supports the acquisition of additional new languages (Dixon et al. 2012; Wagner et al. 2007). Other demographic information is summarized in Table 1.

Instruments

The Motivated Strategies for Learning Questionnaire

The Motivated Strategies for Learning Questionnaire (MSLQ) is an instrument for gauging learners' motivational orientations and learning strategies (Pintrich et al. 1991). The MSLQ was selected as it provides a comprehensive assessment of learner motivations, which is a key component of understanding learner behaviour (Silva 1993). The MSLQ is a reliable instrument, with the reported Cronbach's alpha for its subscales ranging from .52 to .93 (Pintrich et al. 1991). The lowest reliability (.52) is for the help-seeking subscale, which may be because the subscale asks about help-seeking behavior from peers and instructors, though learners may rely on only one source for feedback (Pintrich et al. 1991). The MSLQ is a widely used instrument that can be employed across a range of contexts with reasonable reliability (Taylor 2012). Learners self-report on a Likert scale from 1 (not at all true of me) to 7 (very true of me). The MSLQ contains two scales: motivation and learning strategies, each with multiple sub-components. The motivation scale contains three components: value

Table 1 Participant Demographic Information

ID	Age	Gender	Place of Origin	Languages spoken	Time in Canada	Education Level	Profession
1	35	Female	Syria	Arabic	1 year	Diploma	Education
2	37	Female	Iran	Farsi	Over 10 years	Masters	IT
3	36	Female	Iran	Farsi, French	1 month	Bachelor's	Biochemistry and IT
4	44	Male	Iran	Persian, Azari	2 years	Bachelor's	Engineering
5	58	Male	Brazil	Portuguese	Over 10 years	Bachelor's	Architecture
6	30	Female	China	Mandarin	1 month	Bachelor's	IT
7	49	Male	Syria	Arabic	1 month	Bachelor's	Engineering, Business
8	40	Female	China	Mandarin	3 months	Masters	Engineering
9	33	Female	Iran	Farsi	6 months	Diploma	Tourism
10	36	Female	Iran	Farsi	1 year	Bachelor's	Chemistry
11	59	Female	Iran	Farsi	4 months	PhD	Medical/Healthcare
12	33	Female	Peru	Spanish, Italian	7 months	Bachelor's	Education
13	31	Female	Iran	Farsi	2 years	Bachelor's	Medical/Healthcare
14	32	Male	Iran	Farsi	4 months	Masters	Law
15	49	Female	China	Chinese	1 year	Diploma	Economics

beliefs (achievement goals and task value), expectancy (self-efficacy and control of learning beliefs), and affect (test anxiety). The learning strategies scale measures three strategies: cognitive, meta-cognitive, and resource management. The test anxiety subscale was adapted as participants were not tested in their LINC classes. The modified scale switched references to tests with writing assignments, and it is referred to as the writing anxiety subscale in this paper.

Epistemic Writing Beliefs Questionnaire

The epistemic writing beliefs questionnaire (EWBQ) was administered (Jones 2008) to measure participants' beliefs about learning to write. This instrument was selected as it focuses heavily on higher-order writing features such as organization and thesis development and less on sentence-level skills such as grammar. There are 26 items, each scored on a Likert scale from 1 (no confidence) to 7 (completely confident). The questionnaire comprises three scales: writing behavior, writing tasks, and writing skills. The writing behavior scale measures a learner's approach to writing. "I give up on written assignments before completing them" is one example of an item from this scale. The writing tasks scale measures the confidence a learner has in his or her ability to communicate via writing. Items consist of tasks such as "Write a summary of a long essay that effectively captures the essence of it." The third scale, writing skills, includes items measuring confidence in the technical aspects of writing, such as "Write with concise, clear sentences that 'flow' together." Cronbach's alpha is 0.85, 0.94, and 0.93 for behavior, task, and skills subscales, respectively. (Jones 2008).

Epistemic Emotions Questionnaire

The epistemic emotions questionnaire is a self-report instrument for measuring emotions during knowledge-based activities (Pekrun et al. 2017). On a scale of 1 (not at all) to 5 (very strong), learners report how strongly they feel each of seven emotions: surprise, curiosity, enjoyment, confusion, anxiety, frustration, and boredom in response to information that may conform to or challenge a learner's beliefs. The epistemic emotions questionnaire has a confirmatory fit index of 0.936. This questionnaire was selected to assess the role of epistemic emotions in mature ELLs' learning processes, and their relationships to learner motivations and beliefs.

IELTS Grading Rubric

The IELTS rubric (IELTS 2019) was selected based on recommendations from ELL instructors because it is a fine-grained, standardized rubric for assessing ELL writing. This rubric was selected over those used purely in academic settings as it better assesses a wider range of both academic and non-academic quality indicators. The IELTS is also one of the few tests that can be used to provide evidence of English proficiency for migration purposes. Essays were graded on a scale from 0 to 9, in 0.5 point increments, along four dimensions: task achievement, coherence and cohesion, lexical resource, and grammar. Essays were graded by an independent, hired ELL instructor with expertise in teaching and grading essays written by adult immigrant ELLs. The instructor also provided feedback and identified mistakes per the rubric.

Feedback Classification Scheme

Peer-feedback was first divided into two categories by the independent instructor: low-level and high-level. Low-level feedback was any minor grammatical or vocabulary correction that reviewers made by directly altering the essay itself. Qualitative comments, where reviewers added their own insights were counted as high-level feedback. As an example, correcting a typo would be low-level feedback. Making a comment "there were many problems with the structure of the essay throughout" would be high-level feedback.

Next, all high-level comments were assessed by the independent instructor for accuracy and categorized per Nelson and Schunn's (2009) coding scheme. The scheme identifies several characteristics of feedback, which include type of feedback (is it praise, problem/solution identification, or summarization?), problem/solution scope (does it refer to the paper, or a part of it?), affective language type (is praise used?), problem/solution localization (can the problem be identified?), problem/solution type (is a problem, solution, or both presented?), problem explanation (is the issue explained?), and solution explanation (is reasoning for the solution provided?). This scheme was selected for its fine-grained classification approach that allows for in-depth analysis of feedback.

Procedure

This phase consisted of two, two-hour sessions spaced four to six weeks apart. At the start of the first session, participants signed a consent form. Participants had received a copy of the consent form at least a week prior to study participation, and they were encouraged to go over it with a trusted friend or family member. Following consent, participants completed a demographics questionnaire that collected information about their background, motivation for taking English language classes, employment status, and career goals. Participants also completed the 84-item MSLQ and the EWBQ.

Participants wrote two essays each in response to provided argumentative prompts adapted from IELTS practice prompts (see Table 2). The researcher clarified the prompts and ensured participants understood the task. Essays were written on the computer; blank paper and pens were provided for notetaking. Only two participants used the paper provided. In both cases, participants wrote a few words in their native language with the English translation next to it, which they had looked up on their phones. At the end of this first session, a one-on-one interview with each participant was conducted to reflect on their responses.

Table 2 Essay prompts

Prompt
Some people spend their entire lives in one place. Others move a number of times throughout their lives, looking for a better job, house, community or even climate. Which do you prefer? Staying in one place or moving in search of another place?
We are becoming increasingly dependent on computers. They are used in businesses, hospitals, crime detection and even to fly planes. Is this dependence on computers a good thing or should we be more suspicious of their benefits?

Twelve of the participants returned to the lab approximately four to six weeks after completing the first session. The other three participants (P4, P10 and P14) did not continue with the study due to scheduling constraints. The returning participants took part in the peer-review session which consisted of participants grading and providing essay feedback. When questioned about prior experience with peer-review, all but P7 reported having assessed a peer's writing in their English classes. They all reported it was done informally and without a rubric.

The researcher explained the IELTS rubric to participants. Participants were then asked to provide feedback on the computer and grade three essays that were written by other participants. Since the IELTS rubric uses technical words and is designed for skilled graders, the researcher thoroughly explained the four dimensions to participants. This rubric was selected as the IELTS is written by many immigrants in Canada because it is required for a large number of educational and employment opportunities. Understanding the IELTS requirements, particularly for learners in informal contexts who do not have instructor support, is a highly relevant task. Participants were told the objective of their feedback was to help the writer improve his or her draft. Three essays were reviewed by each participant. Each participant reviewed an essay of high (score = 7), medium (score = 5), or low (score = 3) quality, as determined by the hired instructor.

To explore participant attitudes towards the source of feedback, participants were told they would be shown feedback on both essays from three different sources. They were told the feedback would be labelled as either coming from an instructor, peer, or machine; and it was assumed that the feedback would be correct. However, all feedback was generated by the instructor. The feedback participants were shown could also incorrectly identify errors. This deceit was necessary to ensure that the feedback was of similar quality so that we could explore their attitudes towards the feedback based on who they believed had provided it: a human authority, a human peer, or the software. This aspect of the study was disclosed to participants at the end. Participants were also shown which of the feedback they viewed was correct or incorrect.

The instructor was asked to generate true and false feedback for each essay. True feedback accurately identified a shortcoming in the writing. To generate false feedback, the instructor was asked to incorrectly identify an error in the essay. Across their two essays, participants were shown six pieces of feedback in total. This feedback was comprised of all combinations of feedback source (instructor, peer, machine) and correctness (true, false). Some of the feedback written by the instructor was modified slightly by the researcher to sound more believable in certain cases, while retaining the original content (i.e., replacing more technical language with colloquial words when the text was labelled as peer feedback).

For each piece of feedback, participants completed the epistemic emotions questionnaire. As well, they were asked how strongly they agreed with the feedback on a scale of one (not at all) to seven (very strong). The objective of this activity was to measure whether the perceived source of feedback affected their epistemic response to it, as well as their ability to assess its accuracy.

Findings

Results and implications from the first phase are discussed in this section. This discussion includes the MSLQ and epistemic writing beliefs questionnaire, the

writing samples, and peer-feedback. We focus on these data as understanding a learners' achievement goal orientations, epistemic beliefs and emotions, and cognitive strategies is essential for designing interventions to scaffold learners towards their goals. Since we recruited from a classroom where student groupings are competency-based, the variance in participants' demographic backgrounds is low (See Table 1).

EWBQ and MSLQ Scales

To understand mature ELLs epistemic beliefs, we analyzed each of the three EWBQ subscales. The mean, standard deviation, median, and inter-quartile range (IQR) are shown in Table 3. These scores are comparable to the epistemic beliefs reported by younger students in a study of 170 college students in an introductory English writing class, which found an average score of 4.2 for the behaviours scale and 4.5 for the task and skills scale (Jones 2008). This consistency with prior findings suggests that both groups of learners have strong beliefs in their ability to write.

MSLQ responses were analyzed to understand mature ELLs' motivations and learning strategies. The average scores per scale item and score distribution are shown in Table 4. Overall, participants reported high MSLQ and EWBQ scores. The highest items were the intrinsic goal orientation, task value, control of learning beliefs, elaboration, and organization MSLQ subscales, which leads to our first finding:

> **F1:** Mature Immigrant ELLs have strong interest in mastering course material, even if they find it challenging, and they believe the material is important for them to learn.

This finding is supported by prior work with university level ELLs. A study of Iranian undergraduate ELLs found that learners had similar levels of intrinsic and extrinsic goal orientations (Shafaei and Nejati 2008), suggesting that learners were equally motivated to study English for both personal and practical reasons. Additionally, a study of Thai University ELLs revealed that students in vocational programs were significantly more performance oriented than those in academic streams. (Koul et al. 2009). This suggests that learners in vocational programs have stronger performance orientations. For mature ELLs, the strong intrinsic motivation may be explained by a desire to feel independent and competent, as they did in their previous educational and professional roles before migrating (Mathews-Aydinli 2008), while the extrinsic motivation may be explained by the immediate need to find employment.

Table 3 EWBQ scale averages

Scale	Mean	SD	Median	IQR
Writing behaviours	4.2	0.7	4.1	0.9
Writing skills	4.5	1.5	5.0	1.6
Writing tasks	4.9	1.2	5.1	0.9

Table 4 MSLQ scale averages

Scale	Subscale	Mean	SD	Median	IQR
Motivation scales	Intrinsic goal orientation	5.9	0.6	6.0	0.5
	Extrinsic goal orientation	5.1	1.8	5.7	2.3
	Task value	5.6	1.0	5.8	1.5
	Control of learning beliefs	5.6	0.9	5.8	0.8
	Self-efficacy for learning and performance	3.3	1.3	3.0	2.4
	Writing anxiety	5.1	0.9	5.3	1.3
Learning Strategy Scales	Rehearsal	5.4	0.7	5.5	1.0
	Elaboration	5.6	0.9	5.8	1.3
	Organization	5.9	0.6	5.8	0.8
	Critical thinking	5.4	0.8	5.4	1.0
	Metacognitive self-regulation	5.0	0.7	4.9	0.8
	Time and study environment management	4.9	0.5	4.8	0.6
	Effort regulation	4.0	1.0	3.8	1.0
	Peer learning	5.3	1.1	5.3	1.0
	Help seeking	5.2	0.7	5.3	0.5

Understanding Scale Responses across Groups

To identify different groups of learners within our participant pool, we performed a descriptive analysis of how ELLs' responses varied across the EWBQ and MSLQ scales. Participants were grouped into a low score (score < = 4.5) or high score (score > = 5.5) category. These cut-off scores were selected as the average score for most of the scales was close to 5.0. Table 5 presents the high and low scores in the writing skills and anxiety scales. As is suggested by the table, those with stronger belief in their writing skills may experience lower levels of anxiety, and those who reported weaker writing skills seemed to experience a range of writing anxiety levels. This finding is supported by prior research. In one case, self-confidence was closely related to second language writing anxiety (Matsuda and Gobel 2004). Consistent with the above results, a study of Taiwanese second language learners found that learner beliefs about their writing skills were a stronger predictor of anxiety than actual performance (Cheng et al. 1999) Also, those with lower self-confidence may feel increased anxiety regarding evaluation of their performance (Cheng et al. 1999). The heightened anxiety that mature ELLs may experience in the classroom can be explained through a variety of reasons, some of which were mentioned by learners in the interviews and are discussed next.

P14, who was the only participant to score a seven on the epistemic writing skills subscale, also reported the lowest anxiety. P11, who scored the lowest on the writing skills subscale, had one of the highest anxiety scores. In a follow-up interview, P11's explanation provided context to these scores: "I'm worried about my writing feedback… Because I live in Canada." P11 reported that these concerns led to feelings of frustration when receiving feedback and prompted an urgency to practice more. Unlike P11 who interpreted critical feedback as "punishment", P7, who has lower anxiety,

Table 5 Writing Skills and Writing Anxiety

ID	Writing Skills	Writing Anxiety	Group
11	1.8	5.2	Low
1	3.3	2.8	Low
9	3.3	4.8	Low
5	5.8	3.2	High
7	5.8	2.8	High
14	7.0	1.0	High

views revision as an opportunity for improvement: "Every time when we read something we can change. Even bigger writers when they read their books, they change." These scores and comments suggest that there may be an association between anxiety and learners' approach to writing tasks. Conversely, low epistemic beliefs may inhibit learning strategies, and so, writing development. One explanation for this interpretation are the similarities between the MSLQ metacognitive strategies and control of learning beliefs subscales, as the data in Table 6 suggests.

The few participants who scored low on metacognitive strategies generally scored lower on control of learning beliefs. P8, who falls into the low group, commented on externals factors, such as "different culture background" and "language habit" that cause an individual to make mistakes in English. Despite identifying these translational challenges, P8 does not engage in pre-planning metacognitive activities and prefers to have a complete mental draft before starting to write: "I keep in my head. I don't want to write. I just think a long time." Learners who are less able to regulate their mental processes, as suggested by a tendency to not pre-plan, are less likely to believe their performance is a result of their efforts. However, we do not see a clear division with the control of learning beliefs scores in the high and low groups (Table 6), which may

Table 6 Control of Learning Beliefs and Metacognitive Strategies

ID	Metacognitive Strategies	Control of Learning Beliefs	Group
1	3.6	4.5	Low
15	4.6	3.8	Low
4	4.6	5.8	Low
8	4.7	4	Low
9	4.8	5.5	Low
6	4.9	5.8	Low
10	5	6.5	High
12	5.1	6.5	High
13	5.2	5.8	High
5	5.3	5.8	High
14	5.8	7	High
2	5.9	6.5	High
11	6.4	7	High

indicate that our participants generally had a high internal sense of control when learning to write. Despite this, the scores suggest that most of the mature ELLs would benefit from support for enhancing self-regulatory behaviour, such as metacognitive skills, and increasing their sense of control over the learning process. To promote consistent self-regulatory behaviour, Greene et al. (2010) emphasized the importance of aligning task conditions to trigger productive epistemic beliefs in computer-based learning environments. They recommend that instructors in online environments make the epistemic features of learning activities explicit to learners (Greene et al. 2010). This analysis leads to our next finding:

> **F2:** Learners who are confident in their ability to learn to write also employ productive metacognitive strategies and seek ways to improve their writing.

Next, participants were again categorized per high and low scores, as shown in Tables 7, 8, and 9, to explore how scores vary across the intrinsic goal orientation and other MSLQ subscales.

Within the peer learning and self-efficacy subscales, the few participants who scored low also scored lower than their peers on intrinsic goal orientations. P11, who scored relatively lower on intrinsic motivation, does not consider herself a good writer, stating: "I am not good in my writing." This indicates that goal orientation may be associated with help-seeking behaviour and confidence about writing. In a study of undergraduate ELLs, students with higher motivation were more likely to seek feedback (Waller and Papi 2017). Similar results have been found in studies with undergraduate students, where intrinsic goal orientation was positively associated with self-efficacy while extrinsic goal orientation was negatively associated (Phillips and Gully 1997). These relationships highlight the importance of encouraging intrinsic motivation among ELLs to promote engagement, confidence, and productive learning strategies. Put another way:

> **F3:** Intrinsically motivated mature ELLs are likelier to seek help from instructors and peers.

We then further explored learner motivations for learning to write. For the task value subscale, no participants fell into the low category as none of them scored below a 4.5.

Table 7 Peer Learning and Intrinsic goal orientation

ID	Peer Learning	Intrinsic Goal Orientation	Group
11	2.5	4.3	Low
7	4.0	6.0	Low
5	5.7	5.8	High
6	5.7	6.3	High
13	5.7	6.3	High
12	6.3	6.5	High
1	6.7	5.8	High
14	7.0	7.0	High

Table 8 Self-efficacy and Intrinsic goal orientation

ID	Self-Efficacy	Intrinsic Goal Orientation	Group
11	2.8	4.3	Low
5	5.6	5.8	High
12	5.6	6.5	High
13	5.6	6.3	High
4	5.8	5.5	High
9	5.8	6.0	High
8	5.9	6.0	High
7	6.0	6.0	High
6	6.1	6.3	High
10	6.1	6.0	High
2	6.4	6.3	High
14	6.4	7.0	High

All participants reported a high belief in the importance of learning the material, despite their varying levels of intrinsic motivation. Participants reported high extrinsic goal orientation, while no patterns between participant scores were observed on the other EWBQ and MSLQ subscales. For example, P14 who had low extrinsic goal orientation (score = 2.7) scored lower on self-efficacy (score = 5.0) than P9, who had the highest extrinsic goal orientation (score = 7.0) and higher self-efficacy (score = 5.75).

Learning context can influence the goals learners adopt (Midgley et al. 2001). An explanation for the high average extrinsic orientation scores and lack of association with other EWBQ and MSLQ subscales in this study may be that many of the participants are under pressure to learn these skills to improve their employment prospects. In this setting, adopting a performance orientation may not negatively impact

Table 9 Task Value and Intrinsic goal orientation

ID	Task Value	Intrinsic Goal Orientation	Group
3	5.5	5.0	High
6	5.5	6.3	High
15	5.5	5.8	High
7	5.8	6.0	High
8	5.8	6.0	High
5	6.2	5.8	High
12	6.2	6.5	High
9	6.3	6.0	High
13	6.3	6.3	High
10	6.7	6.0	High
2	6.8	6.3	High
14	7.0	7.0	High

mature ELLs self-efficacy as it might for younger students (Midgley and Urdan 1995). Though mastery goal orientations have been associated with higher levels of motivation and persistence than performance orientations, this difference may not be as pronounced for mature ELLs (Pintrich 2000).

When asked their reasons for taking English classes, most participants had clearly defined goals. For example, P15 needed to score a 6.5 on the IELTS "to go to university and continue my education in [healthcare field] … it need perfect language because this field is related to people and I should speak a lot". P7 was taking ELL classes to improve his job prospects, as he believes employers prefer individuals with "Canadian experience". P7 further explained that "Sometimes they prefer a native speaker or sometimes they prefer younger people."

The above comments add context to the finding that no participant reported a low task value score. These learners may not enjoy the material in their ELL classes, but they recognize the value of the content in helping them reach their goals. Feedback interacts with goal orientation to influence performance, and can be a powerful scaffolding tool (Shute 2008). More specifically, goal-directed, formative feedback which updates learners on progress towards their goals can motivate and direct learners' efforts (Shute 2008). As many of the learners were motivated to master the course material, they are likely to value feedback that provides information on how well they are learning the material. Therefore, numerical grades (summative scores without justification) may not hold much value for adult ELLs. This is summarized in our next finding:

> **F4:** Formative, goal-directed feedback may be more valuable to adult immigrant ELLs than numerical scores.

Next, we analyzed the common writing mistakes made by participants to explore the type of feedback that mature ELLs need to improve their writing, the types of feedback they provide, the type of feedback they value, and how they respond to feedback on their writing.

Common Writing Mistakes

Essay samples were graded to understand the writing challenges faced by mature ELLs. The breakdown of essay scores across the four dimensions of the IELTS rubric is summarized in Table 10. The average score was 5.1 out of 9. As a guide for academic institutions, IELTS states that a score of 5.5 is probably acceptable for training courses that are linguistically less demanding (e.g., catering, fire services) (Settings IELTS entry scores n.d.).

An overview of the most common weaknesses identified across each dimension of the IELTS rubric is provided below:

- Task response: Participants did not appear to accurately conceptualize all components of the task requirements, as most essays did not adequately meet all aspects of the prompt. All but one essay received some variant of the comment "Addresses the task only partially". Only three participants (P3, P7, P11) reported that they engaged in pre-writing activities, such as note-taking. ELLs who engage in mining

Table 10 Essay average scores

Topic	Mean	SD	Median	IQR
Task Response	5.0	1.1	5.0	2.0
Cohesion	5.1	0.9	5.0	1.0
Vocabulary	5.4	1.2	5.0	1.0
Grammar	5.0	1.1	5.0	2.0
Total	5.1	1.0	5.3	1.5

strategies, where they re-read the essay prompt to search for specific information while writing, have improved task response and overall writing quality (Bråten et al. 2011). Promoting such mining strategies can foster reflection and revision. Additionally, these errors may be related to challenges in understanding writing expectations in a new language, and individual differences may also play a role. As said by P15: "for me even in my first language, I sometimes find it difficult to start."

- Cohesion and coherence: Most of the essays had structural problems. One common grader comment was the "lack of progression" in student essays. Ideas were underdeveloped and lacked supporting details. Many essays also incorrectly used cohesive devices. For instance: "But in compare between good things and bad things from computer, I have to say that it is good device and I love it." Several participants reported that they thought in their native language, which led to translational errors when converting their thoughts to writing. P6 explains: "I think in Mandarin so maybe I have some problem, translate the sentence from Mandarin to English. Maybe after translation is not what I want to say." P11 adds to this "I think it in Persian and write it in English. [Instructor] always said to us think English and live English. I do not agree." Participants' tendency to think about their writing in their native language can be leveraged by encouraging pre-writing activities as brainstorming in one's native language leads to better cohesion and coherence in writing (Lally 2000).
- Lexical resource: Generally, essays contained inaccuracies in collocations and expressions that impeded meaning. The errors suggest learners are attempting to convey complex ideas but are unable to construct the appropriate sentence structure. For instance: "We can sleep less hours that our grandparents slept, we spend hours and hours in front of the …" Learners expressed hesitation when using complex words they knew for which they lacked confidence in its proper usage. P6 said: "I maybe I know the word but I don't know how to use it appropriate." P1 pointed to a perceived lack of rules as her challenge in learning vocabulary: "This is 'a' instead of 'u'. So, and I found this really difficult for me, like just because there is no rules. Like there is... sometimes it's the same word, the same pronoun but the meaning is different." When uncertainty poses a challenge in individual writing, collaboration with peers can provide learners with confidence and greater awareness of their writing challenges (Tang 1999).
- Grammar: Common grammatical errors included run-on sentences, a limited range of structures, and complex, unclear sentences. At least one of these issues occurred in all 30 essays. P1 suggested that having example sentences with the word would

give her confidence in modelling her sentence: "You write one word give you the meaning and you write a sentence, I will write a sentence". This kind of modelling has been shown to benefit adult ELLS When adult ELLs work collaboratively, they produce texts with more complex sentences and higher accuracy (Storch 2005).

Peer-Review Feedback

Mature ELLs provided feedback on peers' writing. Across the 30 essays, a total of 316 low-level feedback comments were made. Of these, 198 (63%) were correct. In total, 68 high-level feedback comments were made. Of these, 71% were correct. Most of the feedback was delivered in an affectively neutral tone (83%), and praise was uncommon (21%).

Feedback most commonly addressed the coherence and cohesion of the essay, at a frequency almost double that of the other three rubric dimensions. Most feedback identified a problem (96%), but only a small portion of this feedback offered a solution (23%). This breakdown is shown in Table 11, and is summarized in the next finding:

> **F5:** While mature ELLs provided more frequent low-level feedback, their high-level feedback was more accurate.

Explanations of the problems and solutions were rare, which may be problematic as learners are more likely to implement feedback revisions they understand, and learners are more likely to understand a problem if a solution is provided (Nelson and Schunn 2009). On the advantage of peer-review, P3 points out that learners with similar challenges may be better suited to help identify errors as "foreigner students have problem in some kind of thing. I know for example, the Iranian the 'how' and 'where' is their problem." Training ELL students by providing examples of good feedback leads to more revision-oriented feedback (suggestions and critical analysis) in computer-mediated peer review (Liou and Peng 2009). P15 says: "I want to read some real good ones to compare because yes, a good writer can use some materials I often I know but I can't write out.". This suggests that one way to improve feedback quality could be through coaching mature ELLs to provide explanations in their feedback.

Epistemic Emotions in Response to Feedback

As the epistemic emotions questionnaire data was not normally distributed, a more robust non-parametric test was run (Wilcoxon Signed-rank test for paired data). Of the seven emotions, participants felt moderately more anxious ($Z = 2.61$, $p = 0.008$), bored ($Z = 2.39$, $p = 0.017$), and frustrated ($Z = 2.41$, $p = 0.008$) when presented with fake feedback. No differences were found for the emotions of surprise, curiosity, excitement, and confusion between the true and false feedback. Median and IQR for each of the responses are shown in Table 12.

There was no significant difference in participants' agreement when they were given true or false feedback ($Z = 1.04$, $p = ns$). Participants were split on whether they believed they could identify incorrect peer feedback. Both P7 ("No it would not

Table 11 High-level feedback breakdown and classifications

Feedback attribute	Feedback Categories	Count	Proportion
Correct	Correct	48	0.71
	Incorrect	20	0.29
IELTS Dimension	Task response	13	0.19
	Coherence/cohesion	29	0.43
	Vocabulary	14	0.21
	Grammar	12	0.18
Feedback type	praise	14	0.21
	problem/solution	53	0.78
	summary	1	0.01
Scope	global	45	0.61
	local	23	0.39
Affective language	mitigation-compliment	10	0.02
	mitigation-other	9	0.15
	neutral	49	0.83
Localization	localized	24	0.41
	not localized	35	0.59
Problem/solution	problem	45	0.76
	solution	2	0.03
	both	12	0.20
Problem explanation	absent	49	0.83
	content	10	0.17
Solution explanation	absent	56	0.97
	content	2	0.03

confuse me … I know what I wrote and where I want to focus on.") and P3 ("No I didn't get confused. Because sometimes if I have a problem I know that I have a problem.") explained that they knew their challenge areas and therefore would not be confused by incorrect peer feedback. Some acknowledged that unless peer feedback

Table 12 Median and IQR for responses to feedback accuracy

Response	Feedback Accuracy	Median	IQR
Anxious	True	2.0	1.0
	False	2.0	2.0
Bored	True	1.0	1.0
	False	2.0	2.8
Frustrated	True	1.0	1.0
	False	2.0	2.0
Agreement	True	5.0	3.0
	False	5.0	4.3

was clearly wrong, they may not identify it as incorrect. P15 said: "If the feedback is very clearly it is wrong, maybe I can tell. Sometimes I can't identify". P6 shared a similar thought: "If he didn't know what is correct and the other give him the wrong feedback it may be confused". While trust in peer feedback was mixed, no participant reported distrust in instructor feedback. When asked if he had ever received incorrect instructor feedback, P7 expressed surprise: "I didn't think of that before. I don't know. I think that the instructor will give the right feedback as he should be experienced."

Participants agreed more ($Z = 3.12$, $p = 0.005$) with what they believed to be machine feedback than peer feedback. However, there was no significant difference in how much they agreed with the feedback between the instructor and a peer ($Z = 2.02$, $p = ns$) or the instructor and machine ($Z = 1.23$, $p = ns$). Median and IQR values for the responses are shown in Table 13. As content of feedback may influence acceptance (Nelson and Schunn 2009), our investigations focused on the source of feedback and its perceived level of authority. The content of the feedback was not manipulated in this study as all feedback shown to participants was generated by the instructor and was therefore standardized. This analysis informs our next two findings:

> **F6:** Participants felt significantly more anxious, frustrated, and bored when presented with false feedback (incorrect feedback that was generated by the instructor).
> **F7:** There was no difference detected in mature immigrant ELL acceptance of peer and instructor feedback. However, mature immigrant learners seem accepting and less likely to question automated feedback than peer feedback.

One explanation of the increased negative emotions (see Table 5: Writing Skills and Writing Anxiety) may be that mature immigrant ELLs have strong and generally correct beliefs about what their writing weaknesses and challenges are. Feedback that contradicts these beliefs elicits negative feelings of frustration and anxiety, especially when they do not receive justification for the feedback or understand why they received it. Previous studies have also observed strong beliefs about feedback correctness among ELLs. A study of peer-review found that graduate student ELLs incorporated only 53% of feedback during revision, suggesting that learners selectively decide which comments are accurate (Mendonça and Johnson 1994). In another study, the most common reason given for rejecting peer feedback was that it was believed to be incorrect (Yang et al. 2006).

Participants were significantly more likely to agree with machine feedback than with peer feedback, but not more likely to agree with instructor feedback than machine feedback. One explanation of this may be that learners initially read peer-feedback with

Table 13 Median and IQR for responses to perceived feedback source

Perceived Feedback Source	Median	IQR
Peer	4.0	1.8
Instructor	5.0	2.0
Machine	6.0	1.0

distrust and instructor feedback with trust. However, they are open to automated feedback until it contradicts a strongly held belief. For instance, P15 expressed distrust of automated feedback in response to feedback she believed was wrong: "The best part of my study is grammar. I don't believe I can make so many frequent grammar errors. So if computer can give this kind of assessment, maybe the computer can point mistakes." P1 explains: "I will trust people or machines who I believe... I know they know more than me." If feedback does not align with their self-beliefs, non-acceptance of automated feedback may increase. These reactions suggest that response to feedback is highly context dependant. Prior research on learner perceptions of peer-feedback in undergraduate and graduate classrooms is inconclusive, with some indicating that ELLs are distrustful of peer advice (Murau 1993; S. Zhang 1995), while others suggest that they are just as or more likely to incorporate peer comments than instructor feedback (Mendonça and Johnson 1994; Paulus 1999).

Participants were also more likely to feel confused about feedback they believed came from a peer or instructor, compared to that from a machine. An explanation for this lack of confusion could be that most of the participants were receiving automated feedback for the first time, which may have meant they had no existing expectations for the feedback they received. Their acceptance may be partly due to automation bias, where an overload of manual and automated input distracts a decision-maker and leads to increased acceptance of automated tasks (Parasuraman and Manzey 2010). The finding that mature ELLs are accepting, and possibly even more trusting of automated feedback than peer feedback, could be useful in designing online learning systems. For instance, coaching mature ELLs about qualities of effective feedback could lead to more meaningful peer-review and be supported by a computational agent.

Discussion: Mature ELL Attitudes towards Learning to Write through Peer-Review

The above section presented the first phase of work contributing to the development of a writing support tool for adult immigrant ELLs. Individual differences were measured through the MSLQ and EWBQ and led to **Finding F4**, which suggests unique ELL characteristics, such as low value placed on summative assessment (Section 4.4.2). From this analysis, we present two suggestions to consider in the design of feedback systems. First, adult immigrant ELLs have high intrinsic motivation, which may be prompted by a need to quickly acquire and apply skills (**Finding F1**, Section 4.4.1). This finding of high motivation among older adult ELLs supplements an earlier study which found that as individuals age, belief in their ability to learn strengthens (Schommer 1998). The correlation between age and epistemic beliefs, as well as strong motivation to learn, suggests that mature immigrant learners may be better equipped to independently learn new and complex skills than younger learners. Second, learners' epistemic writing beliefs and goal orientations may shape the cognitive, metacognitive, and rehearsal strategies used when learning to write, such as help-seeking behaviour, which may ultimately impact writing development and performance (**Finding F2**, **F3**, Section 4.4.2).

Learners who held positive epistemic beliefs also employed productive learning strategies, as has been proposed in prior theoretical models (Bråten et al. 2011; K. Muis 2007). This finding has implications for adaptive system design. For instance, learners with low epistemic beliefs may require greater prompting to seek feedback on their

writing. As suggested by Greene et al. (2010), an epistemic tutor in a computer-based learning environment could provide explicit cues about the learning task to activate productive epistemic beliefs. An example of these types of cues would be prompts to compare and contrast the information presented.

Regardless of learners' epistemic beliefs, the provided peer-feedback was generally accurate and more so when the feedback commented on a high-level aspect of the paper (**Finding F5**, Section 4.4.4). In general, participants reacted differently when feedback was not correct, as shown by an increase of negative emotions when faced with fake feedback (**Finding F6**, Section 4.4.5). Feedback from peers elicited greater confusion, suggesting that learners were more willing to question comments they believed came from another learner. However, participating ELLs tended to be very accepting of automated feedback, more so than peer feedback (**Finding F7**, Section 4.4.5). Studies of younger students indicate that they are generally optimistic about automated feedback, similar to our mature ELLs. While university-level students (both native English speakers and ELLs) appreciate feedback on writing, they are not comfortable with the numerical grades provided by these systems (Curran et al. 2013; Dikli and Bleyle 2014; Fang 2010). However, grades were not a major concern for our learners who were engaged in an informal learning activity (**Finding F4**, Section 4.4.2), which suggests that mature ELLs may be more accepting of automated support than younger students. Thus, a platform incorporating automated feedback may be more likely to be adopted by mature ELLs as positive perceptions of automated feedback are linked to future intention to use the writing support tool (Roscoe et al. 2017).

Many of the participants in this study were no longer taking English classes by the second session. Most were either searching for a job or attempting to return to school. Thus, this study discovered a need for a tool to facilitate writing development outside of the classroom and without instructor involvement. As summarized by P8: "Feedback is important. It's key. Practice. I can practice in my home but we can't if we don't have feedback. Practice will not continue because I don't know if it's right or wrong". In this context, two ways learners can get timely and accurate feedback is through peer-review and machine generated metrics. This analysis found support that peer-review between mature immigrant ELLs can yield high-quality, accurate feedback. Additionally, mature ELLs were generally accepting of machine feedback; a question that has not been previously explored. So, an application that incorporates peer and automated feedback could bridge instructor-led learning with learning outside the classroom, and thus serve as a valuable tool for new immigrants transitioning into the workplace, school, and new culture overall. These kinds of transitional technologies have been found to improve learning outcomes with other learner groups, such as low-literacy adults (Munteanu et al. 2010). To further develop this idea, a more thorough examination of the design requirements for such a tool was undertaken in the second phase.

Phase Two: Developing Design Guidelines

The methods used in education research for early specification of design requirements are still generally grounded in prior knowledge, theoretical frameworks, and the existing body of practice. Often, end users of an application are not involved in its development. While traditional methodologies, such as these, provide a strong

background for development, they do not always easily adapt to a wide range of user needs. However, participatory design does, and this design method helps ensure ease of use. As such, participatory design methods were used to investigate how learners respond to various feedback types and presentations, and to elicit design requirements for learning analytics and feedback features in the final part of this study. The first phase found support for the development of a platform that allows mature immigrant ELLs to engage in peer-review, while this second phase aims to provide further context to the earlier findings.

Methodology: Participatory Design

Participatory design (PD) integrates users into the technology creation process through a variety of methods (e.g., interviews, observations, or design activities) that elicit requirements from the early stages of the design process. A key component of PD is the partnership between the researcher and participant, where the researcher acts as apprentice to the participant who is a master of their process (Soegaard and Friis Dam 2013). PD can be employed with a variety of users, even if they do not have any design experience or technical knowledge, and it can reveal hidden elements that result from the difficulty associated with verbalizing one's process. Though it is extensively employed in other research domains, PD is rarely used in educational contexts (Birch and Demmans Epp 2015) despite its potential to inform design by better understanding learners and their environments.

Combining observations from PD with traditional assessments can provide additional insight. One major advantage of integrating these approaches is that it provides both an objective view of the learning context and the learner's perception of it. This can highlight surprising (in) consistencies between the two (Liaqat et al. 2018). Another advantage of PD is that it can help generate design ideas. The result of such investigations are themes and recommendations, which are followed by more qualitative research that aims to formalize the designs of technologies for the target user group. This is the approach we took in the second part of our research. These qualitative design-based methods are easily combined with more quantitative insights, such as those drawn from psychometrics and learning analytics as was done in the first phase of our research.

Procedure

Two open-ended PD group sessions were run where ELLs were given tasks that would support the design of a tool. Both groups were facilitated by the same researcher. A subset of the previous participants were included in these PD group sessions. Each session lasted approximately 1.5 to 2 h. The first group consisted of three participants (P1, P7, P12), while the second group involved four (P3, P8, P11, P13). These groups were formed based on participant availability. See Table 1 for participant demographic information.

For both PD sessions, participants first completed a warm-up activity where they were asked to work together and provide feedback on a drawing to get them comfortable with the type of activities performed in the main session. The main session then varied by group.

The first PD group was tasked with modifying the IELTS rubric into one they would feel comfortable using for peer-review, as it was an assessment tool all the participants had experienced using. They were provided with pen and paper to sketch out their ideas. The group was encouraged to discuss their ideas and to reach a consensus on design decisions. This activity was designed so that the think aloud process and group collaboration could reveal how the target learners translated a set of standardized requirements into their own words. This activity was also meant to help identify gaps between the learners' interpretations and the expectations communicated through the rubric.

The activities of the second group were independent from those of the first. The group was asked to collaboratively peer-review an essay. They were asked to think aloud and to reach a consensus with the group before providing feedback. The observed workflow was then compared with the modified rubric developed by the first group to identify common themes across the two. Since participants were at LINC level two, all had experience in assessing writing (whether their own or a peer's). They also had experience with receiving feedback. This activity was designed to elicit information about what learners emphasized and valued when providing feedback.

Themes were drawn from the transcribed sessions using thematic analysis methods (Braun & Clarke, 2006). Thematic analysis makes use of inductive strategies for analyzing qualitative data, where data is iteratively abstracted to identify common themes and patterns. This inductive approach allows for a systematic analysis of qualitative data and links research with theory (Smith, Harre, & Langenhove, 1995). We report structured findings as common for mixed methods research where thematic analysis is required. Similar to other work with ELLs, we provide support for the findings (emerging themes) through the use of quotes from participants (Demmans Epp 2017; Munteanu et al. 2010).

Findings

Themes and related codes were generated by the researcher from each of the two PD group sessions, as shown in Fig. 1.

Explicit Cues Are Needed to Scaffold the Writing Process

Rubrics Need to Embed Guidance to Support Peer-Reviewers The first group's preliminary design included a slightly reworked version of the IELTS rubric,

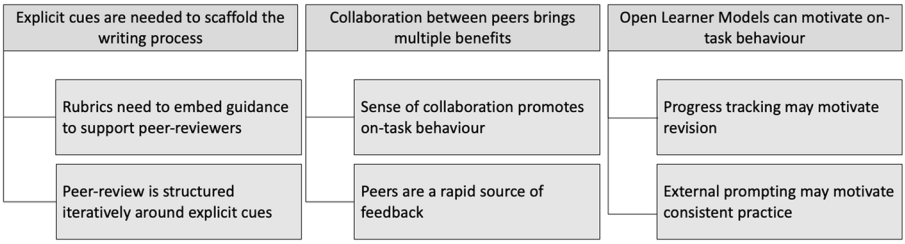

Fig. 1 Themes and related codes

where the peer-reviewer could give a numerical score and provide open-ended comments under each dimension. Overwhelmingly, the group decided they did not like the open-ended idea of an IELTS rubric, with P1 (Table 1) explaining they did not feel they had the expertise to provide detailed feedback for each dimension: *"For me, it's too complicated, and I can't read exactly the difference between each category."* Other suggestions were made, such as having a guiding set of questions that followed the structure of the essay, as said by P12: *"We could follow the steps of the paragraph. We could start with the topic sentence, then work on a specific topic, give an explanation for each one, and then the conclusion."* The group concluded that it was important to get a holistic view and to have the ability to see the details of the writing.

Peer-Review Is Structured Iteratively around Explicit Cues The second group was asked to review an essay together. Without being prompted to do so by the researcher, participants made sure that everyone in the group agreed with a point before writing it down. When asked to develop a rubric, there were conflicting ideas on how to organize it. The participants agreed that the first criteria should be the structure of the essay: *"I think first there should be an introduction. We need the question but in different words and our main idea that we want to analyze it in whole essay"* [P13].

The group read through the essay at least three times. After the first pass, the group discussed the structure and task response of the essay. P13 remarked: *"It didn't show us any point or goals of the writing. There's no introduction, no conclusion, just examples."* Next, the group agreed to move on to check for vocabulary, grammar, and punctuation. However, they realized they would have to re-read the essay to observe these issues as they had been focusing on structural aspects the first time. In the last pass, the group looked at more holistic aspects of the essay, such as the writing style and overall impact, as explained by P8: *"In Chinese, I'm focused on some writing skill about how you organize your sentence, how to make the idea clearly, and make the reader want to read your writing."*

The first PD group discussed how using a purely holistic rubric, such as the IELTS was a challenge for them. The second group generated a deconstructed rubric that transformed review into a linear process. Rather than attempt to read through the entire essay before being able to use the rubric, the group could comment on each part of the essay as they read it. Additionally, by focusing on a distinct aspect of the writing in each pass, it seemed that less cognitive effort was required overall.

Collaboration between Peers Brings Multiple Benefits

Sense of Collaboration Promotes on-Task Behaviour While developing the modified rubric, the first group discussed how it would be helpful if they could see what feedback other reviewers had given. The group appreciated the idea of having a sense of collaboration with other learners. Another benefit mentioned was that it may draw attention to overlooked problems. As mentioned by P7, *"It will help you pay more attention to the details."* P12 confirmed that *"it keeps good information between us, keeps us posted."*

On the benefits of being able to ask questions, P1 said:

> That's a good idea. That would keep you writing more, because you write more, by asking questions. Because the people that give you the grade, you are going to ask, oh why, what is the mistake, so you write more, the relation between you and the other people.

There was consensus in the group on the importance of two-way communication channels between peers. The group believed that communication fosters a sense of collaboration, motivates, and helps learners identify and understand their errors. In other contexts, the benefits of a two-way, semi-structured communication platform have included timely responses to queries and higher quality of help (Greer et al. 1998a, 1998b) as well as increased social support that contributes to learning (Phirangee et al. 2016).

Peers Are a Rapid Source of Feedback The reviewing behaviour in the second group demonstrated additional advantages of collaboration. When a participant went off-track during the review, another participant would bring them back on task. At one point, P11 began discussing what their own response to the prompt would be. The participant was quickly reminded about the task objective by P13: *"Ok, it's right, but we just want the feedback about this essay, not our opinion."* Participants were also comfortable asking each other for clarification when they encountered unfamiliar words or sentence structures while reviewing. For instance, P11 asked *"What's the meaning 'wished' in this sentence?"*, to which P13 replied *"It's used a verb, but in the past"*. Learners may also be more willing to seek help from peers. With PHelpS, a peer-help system for the workplace, employees asked questions frequently to other co-workers when provided with a platform that facilitated this exchange (Julita Vassileva et al. 2016).

Collaboration allowed the groups to identify opportunities for improvement and also promoted on-task behaviour. Through discussion, learners could seek immediate clarification from peers, remind each other when discussion went off-track, and collaboratively make sense of the learning activity. These interactions demonstrate the important role peers play in regulating learning behaviour among mature ELLs, suggesting that providing opportunities for collaboration in an online platform could bring similar benefits.

Open Learner Models Can Promote on-Task Behavior

Progress Tracking May Motivate Revision The first group discussed how they found it a challenge to review and revise their writing. They suggested that prompts to revise, such as a rubric for self-assessment might motivate them to review their writing. As P1 stated: *"I don't like to review my writing, so maybe now, let's see how many mistakes I can find"* with P12 adding *"I like seeing numbers."* P1 illustrated that incorporating a means of assessing performance could be a strong incentive when faced with a lack of intrinsic motivation, while P13 emphasized the importance of concrete indicators of performance. In language learning settings, the use of open learner models to provide feedback on current learner abilities and progress has been found to support continued

learner effort (Demmans Epp and McCalla 2011). Beyond motivating learning activities, prompting learners to reflect on models that represent their current knowledge or skill state can facilitate knowledge acquisition (Zapata-Rivera and Greer 2002). OLMs can also encourage cognitive processes, such as noticing and depth of processing, that are expected to support language acquisition (Shahrour and Bull 2008).

External Prompting May Motivate Consistent Practice The second group explained how they need better tools to help them learn new vocabulary and grammar rules. Simply saving new information to a library was not enough: [P13] "I think it's not good enough. It's just a collection of words." [P3] "I know some applications help you to repeat again and again. If it's just a collection you put it aside." Consistent with prior findings of migrant ELL use of technology (Demmans Epp 2017), the group agreed that it was insufficient to just save. They wanted a system that would prompt them to regularly practice what they had saved in their library, which is also reflective of their need for structure when writing.

A common theme in mature immigrant ELL discussion seems to be an aversion to uncertainty and vague objectives when learning to write. Providing more tangible, quantifiable details, better explanations, or mechanisms that guide learners through an activity seem to be potentially useful scaffolds for supporting learners during the writing process. This can be done through learning dashboards or open learner models which have been found to increase motivation and engagement for ELLs (Tsourounis and Demmans Epp 2016). A simple feature addition that could meet these learner preferences would be the use of an OLM that allows learners to reflect on their performance, plan their future learning activities, or see their progress (Demmans Epp and Bull 2015).

Discussion: Co-Designing Peer-Review for Mature ELL Needs Participants emphasized the need for structured direction throughout the writing process. This need for explicit guidance may be explained by the phase one findings, which revealed a strong association between epistemic writing beliefs and learning strategies, as well as links between anxiety and time and study environment management. This suggests that when learners lack confidence in their writing ability, they are less likely to engage in effective learning strategies, and so feel a greater need for clear structure.

Like other adult migrant ELLs (Demmans Epp 2017), participating ELLs repeatedly highlighted the importance of socio-collaborative approaches to learning throughout the PD process. This type of approach is supported by language-learning theories, such as languaging, that state language is learned through deep interaction with that language and those in our surroundings (Swain 1995). In the first phase it was found that learners have high intrinsic motivation to learn to write, and they can provide high-level peer feedback with 63% accuracy. Combined, these findings add support for the need to build an online peer-review platform for writing development. Such a platform could even integrate just-in-time support from peers that can be used to address specific writing problems as identified through the feedback that a learner has received. This type of interaction would be consistent with the aims of the PHelpS and iHelp systems that were previously used to support learners who were trying to write code and adults in the workplace (Bull et al. 2001; Greer et al. 1998a, 1998b; Julita Vassileva et al. 2016).

Peer collaboration seems to have many benefits. A study of collaborative writing among adult ELLs found that learners working in pairs produced shorter texts than individual writers, but these texts had higher complexity, grammatical accuracy, and better responded to the task (Storch 2005). Additionally, computer-mediated collaborative writing provides a platform for learners to engage in meta-discussions about language and writing, which may promote deeper levels of understanding of the material (Storch 2011; Swain 1995). Thus, a peer-network may be an efficient and effective channel for providing learners with timely and useful writing feedback. Such peer-networks have shown potential in other domains for providing just-in-time, meaningful support (Bull et al. 2001; Julita Vassileva et al. 2016). They have also resulted in better second-language writing products (Abrams 2019).

The PD group sessions revealed that members of this population may lack intrinsic motivation to complete certain writing tasks, such as revision. Additionally, mature immigrant ELLs' discussion indicates they appreciate having clearly defined milestones and seeing their progress towards them. Thus, there is a need to more explicitly link in-app achievements to real-world, tangible progress. This can be done through open learner models that promote reflection and allow learners to contextualize progress through comparison with their prior abilities or peers, as has been seen in other language-learning contexts (Demmans Epp and McCalla 2011). Providing more tangible, quantifiable details, better explanations, or mechanisms that guide learners through an activity seem to be potentially useful scaffolds for supporting learners during the writing process.

Implications and Design Guidelines

The cyclical writing process of submitting a draft, receiving feedback, and revising can take considerable class time and is resource intensive for instructors. It also does not support the transition of learners into the workplace environment, which is particularly challenging for mature migrant ELLs as they have limited access to formal support systems when transitioning to this context. Computer-based support tools provide opportunities to guide learners through the writing process by provisioning timely feedback and supporting the recursive nature of writing that traditional classrooms often do not support. In addition to the adaptive supports and automated assessment one would expect in such tools, integrating online peer review could be beneficial because it can lead to greater on-task behaviour than face-to-face interaction (DiGiovanni and Nagaswami 2001). From the data captured in the PD group sessions and by drawing on our earlier findings, we suggest four design guidelines for the development of writing support tools for mature ELLs. These design guidelines and example implementations are summarized in Table 14. In this section, we detail these design guidelines and, highlight their connections to existing literature where appropriate.

Design Guideline One: Support Collaborative Writing Processes

The study found that learners were able to provide accurate feedback, especially on higher-level issues such as task response and logic (**Finding F5**, Section 4.4.4). As

Table 14 Summary of Design Guidelines

Design Guidelines	Example Implementation
DG1: Support Collaborative Writing Processes	Encourage learners to discuss concerns in a group channel (e.g. Are you unsure about this feedback? Ask your peers for their opinions in the main discussion channel.)
DG2: Enhance Peer Support with Automated Support	Adaptive prompts to encourage learners to reflect on task requirements as a pre-writing activity (e.g. What do you think the prompt is asking for when it says *X*? How could you address that?)
DG3: Structure Tasks to Support Linear and Holistic Review	Ask peer-reviewers localized questions (e.g., Does this sentence support the topic statement?) and holistic questions (e.g. What are your overall impressions?).
DG4: Contextualize Learning through an Open Learner Model	Connect in-system performance to practical skill acquisition (e.g., "You can now write short, work-related emails to a co-worker")

part of meeting this objective, the workflow of peer-review was observed to understand how a system might support this process. It was discovered that as peer-reviewers, mature immigrant ELLs felt comfortable collaborating, asking questions, and correcting one another. Mature ELLs' tendency to question peer feedback also suggests that peer-review can serve as a platform for productive discussion (**Finding F7**, Section 4.4.5). This finding is supported by a study which suggests that face-to-face peer-review can serve as a scaffolding mechanism for ELLs (Guerrero and Villamil 2002). However, group dynamics (e.g., relationship between peers, cultural background, type of task) affect learners' engagement with online peer-review, highlighting the importance of properly designing the mechanisms for delivering peer-review systems (Yu and Lee 2016). Otherwise, mature ELLs may not trust the system or engage with it.

We also found that mature ELLs feel anxious when presented with feedback that is inaccurate (**Finding F6**, Section 4.4.5). Providing a channel where learners can question the feedback they receive may help alleviate some of this anxiety. This can be done in a manner that is consistent with the approaches used in open learner models, where learners can try to persuade the system (Bull et al. 2016), edit the results of an assessment (Mabbott and Bull 2006), provide additional evidence of their knowledge or skills (Bull and McKay 2004), or negotiate with the OLM (Bull and Pain 1995; Kerly et al. 2007; Mabbott and Bull 2006) when they disagree with something that is communicated through that OLM.

In existing writing systems, the peer-review process is rigidly structured. Peer-reviewers communicate with the writers through a rubric, and writers usually cannot respond back. This study found support for creating a platform that allows for more free-form communication and discussion between learners, and that helps reviewers provide more formative feedback (**Finding F4**, Section 4.4.2). Additionally, our learners were intrinsically motivated (**Finding F1**, Section 4.4.1), which suggests they are more likely to engage in help-seeking behaviour (**Finding F3**, Section 4.4.2). These findings lead to our first design guideline:

DG1: Peer-review platforms should allow mature immigrant ELLs to collaboratively iterate through the writing process.

Design Guideline Two: Enhance Peer Support with Automated Support

The acceptance of peer-review feedback was also measured, which showed that mature ELLs are more accepting of machine feedback than peer feedback (**Finding F7**, Section 4.4.5). This suggests that automated tutors (i.e., agents) could play an important role in prompting and guiding learners through writing tasks. We propose three types of support that could be provided by an online tutor, based on the patterns of errors learners made in each of the four IELTS rubric dimensions. This includes supports to help learners conceptualize, structure, and include details in their writing.

Conceptualize The first challenge participants faced was interpreting task requirements. One strategy for addressing this issue is pre-writing, an essential first step in the writing process for idea development (Rohman 1965). It seems that participants were not familiar with pre-writing, as the majority did not use the provided paper for notetaking. Writing a multiple paragraph essay without an outline can lead to incomplete, off-topic responses and exacerbate structural issues. So, a support tool should ensure that learners understand the task requirements. This can be done by using an agent or other mechanism based on a user model to ask leading questions that prompt learners to explicitly define the task as a pre-writing activity. For instance, the iHelp system matched learners with a peer based on a model of both students' knowledge of domain content (Julita Vassileva et al. 1999). A similar approach could be used to match adult ELLs to ensure that the peer providing feedback has the knowledge necessary to support the writer whose work is being reviewed.

Other pre-writing activities often used in classrooms to help learners make sense of the task requirements are clustering, where learners brainstorm related words or thoughts in response to a prompt, rapid free writing, where learners quickly write down ideas, and asking wh-questions where learners come up with who, what, where, when, why, and how questions (Richards and Renandya 2002). Online tools could further enhance these activities by serving as a platform for learners to springboard ideas and seek validation and feedback from peers.

Structure It was clear from many of the essays that their authors struggled with framing coherent arguments. The lack of high-level structure, such as paragraphs, suggests that not all learners were familiar with the essay format. Teaching structure is a challenging task as practice may not always improve skill. This may be because learners require a conceptual understanding before being able to transfer structural features to other contexts (Richards and Renandya 2002). Support tools can help by providing templates that learners can select and build from. On a more localized level, structural issues with sentences were also common, and similar templates could be provided.

Details Many essays did not provide an adequate level of detail. Feedback mechanisms could prompt learners to develop their points as they write by suggesting sentence starters that encourage writers to provide different types of details, by including an example, elaborating, or providing a counterargument.

In addition to these supports, mature ELLS have a need for greater guidance on technical aspects of writing. It was found that the accuracy of the low-level feedback reviewers provided was lower than that of the high-level, suggesting that mature ELLs have incorrect understandings of such technical details which require revision (**Finding F5**, Section 4.4.4). Mature ELLs also need guidance in providing higher-level feedback, which suggests that they may not be aware of such issues in writing and need greater guidance (**Finding F5**, Section 4.4.4). Peer-review is one channel for such guidance. Natural language processing approaches are another if they are augmented with appropriate explanations for different types of errors (Wanderley and Demmans Epp 2020).

Migrant ELLs' tendency to question peer feedback (**Finding F7**, Section 4.4.5) and their strong epistemic beliefs suggest that peer-review can serve as a productive platform for promoting discussion, seeking support, and making sense of feedback together. Peer-review could encourage learners to question and revise their beliefs. Learning can also be facilitated through prompting learners to justify the claims they make about their knowledge (Zapata-Rivera and Greer 2002). However, mature ELLs' higher acceptance of machine feedback suggests that automated guidance can play an important role in supporting technical aspects of the writing process. Additionally, this support would serve as another channel for receiving formative feedback that helps mature ELLs identify and revise errors, which our learners indicated was more valuable to them than numerical scores (**Finding F4**, Section 4.4.2). These findings inform our next guideline:

> **DG2**: Automated tutors should enhance a peer-review platform by scaffolding mature immigrant ELLs through technical aspects of writing by providing prompts for understanding task requirements, improving structure, and providing relevant details.

Design Guideline Three: Structure Tasks to Support Linear and Holistic Review

Learners reviewed writing iteratively, focusing on one aspect of the writing at a time. First, they examined the structure to ensure the required elements of the writing were present. Here, they also assessed the task response. Next, they moved on to correcting the grammar and punctuation. Finally, they engaged in a holistic assessment of writing style and clarity.

Rubrics are generally designed for holistic assessment and assessment is usually done as a last step in the peer-review process. However, mature ELLs assessed writing in a linear pattern as well as a holistic one. This observation suggests that deconstructing the task of peer-review into smaller ones, prompting learners to assess as they read, and encouraging them to take multiple passes may make the peer-review process more manageable.

Research on differences in feedback type in online and face-to-face peer-review is mixed, with some suggesting that an online platform prompts greater global-level feedback, while other studies have found the opposite (Yu and Lee 2016). This may be due in part to individual and context differences between the studies. Understanding

these differences is needed to guide the design of scaffolds in online peer review. It was found earlier that mature ELLs provided high-level feedback with less frequency than lower-level feedback (**Finding F5**, Section 4.4.4), though they prefer to receive high-level formative guidance on their own writing (**Finding F4**, Section 4.4.2). Deconstruction could help address this issue by guiding the peer-reviewer's attention through the different levels of granularity they should assess writing on. So, our next guideline is:

> **DG3**: Rubrics should deconstruct peer-review into manageable tasks that allow mature immigrant ELLs to review both linearly and holistically.

Design Guideline Four: Contextualize Learning through an Open Learner Model

Another key component for this system is determining the kinds of prompts that would motivate learners to engage in the writing process. One finding was that mature immigrant ELLs enjoy assessing themselves, though they may not be sure how to do so. This finding was in line with the earlier results from psychometric scales, which showed that these learners have strong intrinsic motivation to master the material (**Finding F1**, Section 4.4.1), and a strong desire for formative feedback that helps them in assessing progress (**Finding F4**, Section 4.4.2).

A major concern participating ELLs expressed was the danger of this system becoming a tool for collecting vocabulary and writing feedback without it supporting learning and improvement. To address this concern, participating ELLs suggested that regularly receiving prompts to review would help them stay on track, which complements the earlier finding that intrinsically motivated learners are more likely to engage in help-seeking behaviours (**Finding F3**, Section 4.4.2). These prompts for activities, combined with regular feedback and writing tips may help motivate mature immigrant ELLs to consistently practice and improve their writing. Explicit prompts and immediate feedback can be used to promote on-task behaviour (Deterding et al. 2011), such as motivating learners by allowing them to compare progress with each other, with themselves, or with a model of an expert (Demmans Epp and Bull 2015; Domínguez et al. 2013).

While mature ELLs may be motivated through competitiveness with their peers, they also tend to have high intrinsic motivation and need to make sense of their own performance. Coupled with their desire for external prompting, allowing learners to benchmark their own progress, as is common in OLMs, may be a strong motivator. A system that explicitly links a learner's writing performance to the real-world contexts in which those skills are applicable could motivate learners to reach the next benchmark. For instance, relating in-system performance to practical skill acquisition (e.g., "You can now write short, work-related emails to a co-worker") may be more meaningful for mature ELLs than a self-contained feedback system that only makes sense within the platform. The explicit links to real world progress may also help counter the anxiety mature ELLs experience when they perceive feedback to be inaccurate, as such information could help learners make sense of the reasoning behind that feedback (**Finding F6**, Section 4.4.5). These observations lead to the final guideline:

DG4: Open learner models that contextualize ELL progress in the real world should be incorporated into peer-review platforms to motivate mature immigrant ELLs to consistently improve their writing.

Conclusion

Mature immigrant ELLs take English classes for very brief periods of time, often between the time when they arrive in a new country and when they intensively start their job search. During this gap between leaving classes and finding employment, they have no access to formal instruction for improving their writing skills. Additionally, mature ELLs face challenges specific to their age, such as slower acquisition of language comparative to younger ELLs (Johnson and Newport 1989). As such, these learners need a system that provides them with regular opportunities to practice their writing and one that has mechanisms for feedback and support to allow learners to improve. These systems must understand the needs and goals of these learners to support them in attaining their goals (Orji et al. 2018). This study sought to determine what the design requirements of this system should be. The design requirements for this system come from three sources: analysis of the learners' various psychometric constructs, writing samples to identify patterns of errors, and group sessions incorporating participatory-design methods.

Through questionnaires, interviews, and design sessions, we generated seven main findings. We identified that mature immigrants have both high intrinsic motivation and a strong belief in their ability to learn to write, which reinforced the importance of instructors making the epistemic features of a learning task explicit (**Finding F1**, Section 4.4.1). We highlight the need for a tool that better supports the highly independent, self-regulatory behaviour of mature immigrant ELLs (**Finding F2**, **Finding F3**) by providing formative feedback rather than numerical scores (**Finding F4**, Section 4.4.2). This formative feedback is especially necessary for higher-order features because most learners struggle with higher-order features, such as cohesion and structure, when learning to write. However, they can accurately identify problems with higher-order features in their peers' writing (**Finding F5**, Section 4.4.4). This peer support may be more valuable when there are knowledge differences in a group which allow the pairing of learners with complementary skills (Greer et al. 2001). Automated feedback can supplement this peer support by providing critical feedback to learners. We found that receiving incorrect feedback elicits negative emotions in mature ELLs (**Finding F6**, Section 4.4.5), but they were less likely to question automated feedback (**Finding F7**, Section 4.4.5). This automated feedback may also help them to overcome the reported inability of existing tools to facilitate consistent practice and to motivate change in learning strategies and attitudes (Orji et al. 2019).

Building on the above findings, we developed four guidelines for a system to support ELL writing development. First, systems should enable peer support as mature ELLs may benefit from collaboration with peers because they are comfortable questioning one another and can provide accurate, high-level feedback (DG1). Second, since learners are accepting of automated feedback, such feedback can scaffold learners through technical skills involved in writing tasks (DG2). Third, mature ELLs expressed a need for greater direction and were found to review writing iteratively. Thus,

embedding guidance into rubrics and clearly defining a task structure would help learners better understand expectations and reduce anxiety (DG3). Finally, incorporating open learner models may promote reflection, help contextualize performance through comparison with an appropriate frame of reference, and support the mapping of in-app performance to real world skills, thus motivating consistent practice (DG4).

The idea of using learners as peer helpers is a pervasive one within the AIED community (Bull et al. 2001; Julita Vassileva et al. 1999, 2016). The key to such systems is personalizing technology so that learners can find the right helper at the right times (Julita Vassileva et al. 2016). We would argue that this can be taken further by augmenting peer support through automated analytics and their presentation via an open learner model, as is commonly done in intelligent tutoring systems (Demmans Epp and Bull 2015; Zapata-Rivera and Greer 2002). These types of feedback have also been provided in collaborative online learning environments to support student motivation and participation (Bull et al. 2001; J. Vassileva 2008). Regardless of the environment in which these analytics are used, their visualization can support the decision-making of those who are planning learning experiences (Brooks et al. 2014), as is the case with adult migrant ELLs. The findings from this research build upon many of the themes that were woven throughout Jim Greer's work. This contribution includes the expansion of peer support to a population whose needs have been underserved, and the integration of open learner models within this peer-support process. This expansion is accompanied by design guidelines for personalizing the learning experience of mature immigrant ELLs. A peer-support platform that incorporates these design elements would be particularly valuable when mature immigrant ELLs are left without resources, structure, or guidance in the overlooked, and sometimes, long, gap between when they conclude their classes and find employment.

Building on Jim Greer's Legacy

In this research, we extend Jim Greer's work to a domain outside of which he typically worked (mature, migrant ELLs). We build on the various foci of his work, which are evident throughout this and our other research, from the use of analytics and open learner models to the use of peer support and integration of technology into classroom and other learning environments. Here, we have used analytics and feedback to study learner responses to this feedback alongside a form of peer support. While our students were not directly seeking support for knowledge gaps that they had noticed in themselves, they were wanting additional support with respect to improving their writing. To address this need for support, we used the practical approach of engaging peers in this help process. This approach is characteristic of what we view as Jim's highly pragmatic view of educational technologies. He would find simple solutions that would combine some element of artificial intelligence (e.g., learner modelling) or accompanying technology, with a human-in-the loop approach (e.g., peer help, open learner models) to support students in real learning environments. He emphasized getting support out to learners so that it could help them over the importance of using advanced algorithms. This emphasis on impacting practice is one of the many lessons that we will carry with us as we navigate the complex socio-technical systems that we are trying to create by integrating advanced technologies into existing learning environments and systems.

References

Abrams, Z. (2019). Collaborative writing and text quality in Google docs. *Language Learning & Technology, 23*(2), 22–42. https://doi.org/10.1254/4681.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271. https://doi.org/10.1037/0022-0663.84.3.261.

Baecker, R. M., Grudin, J., & Buxton, W. A. S. (1995). In S. Greenberg (Ed.), *Human-computer interaction: Toward the year 2000*. Morgan Kaufmann Publishers Inc..

Birch, H., & Demmans Epp, C. (2015). Participatory design with music students: Empowering children to develop instructional technology. *American Educational Research Association Annual Meeting*.

Bodily, R., Kay, J., Aleven, V., et al (2018). Open learner models and learning analytics dashboards: a systematic review. In: Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK). ACM Press, pp 41–50

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48–70. https://doi.org/10.1080/00461520.2011.538647.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063o

Brooks, C., Winter, M., Greer, J., & McCalla, G. (2004). *The massive user Modelling system (MUMS)*, *In J. C. Lester, R. M. Vicari, & F. Paraguaçu (Eds.), Intelligent Tutoring Systems (Vol. 3220, pp. 635–645)*. Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-540-30139-4_60.

Brooks, C., Greer, J., & Gutwin, C. (2014). The data-assisted approach to Building intelligent technology-enhanced learning environments. In J. A. Larusson & B. White (Eds.), *Learning Analytics (pp. 123–156)*. New York: Springer. https://doi.org/10.1007/978-1-4614-3305-7_7.

Bull, S. (2020). There are open learner models about! *IEEE Transactions on Learning Technologies*, 1–1. https://doi.org/10.1109/TLT.2020.2978473.

Bull, S., & Kay, J. (2010). Open learner models. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Advances in Intelligent Tutoring Systems (pp. 301–322)*. Springer. https://doi.org/10.1007/978-3-642-14363-2_15.

Bull, S., & Kay, J. (2016). SMILI☺: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education, 26*(1), 293–331. https://doi.org/10.1007/s40593-015-0090-8.

Bull, S., & McKay, M. (2004). An open learner model for children and teachers: Inspecting knowledge level of individuals and peers. In J. C. Lester, R. M. Vicari, & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems (Vol. 3220, pp. 646–655)*. Springer. https://doi.org/10.1007/978-3-540-30139-4_61.

Bull, S., & Pain, H. (1995). "Did I say what I think I said, and do you agree with me?" inspecting and questioning the student model. *Proceedings of the World Conference on Artificial Intelligence in Education*, *501–508*.

Bull, S., Greer, J., McCalla, G., & Kettel, L. (2001). Help-seeking in an asynchronous help forum. Proceedings of workshop on help provision and help seeking in interactive learning environments, International Conference on Artificial Intelligence in Education. http://users.sussex.ac.uk/~bend/aied2001/bull.pdf

Bull, S., Ginon, B., Boscolo, C., & Johnson, M. (2016). Introduction of learning visualisations and metacognitive support in a persuadable open learner model. In*Proceedings of the sixth international conference on Learning Analytics & Knowledge - LAK '16* (pp. 30–39). https://doi.org/10.1145/2883851.2883853.

Caine, K. (2016). Local standards for sample size at CHI. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 981–992. https://doi.org/10.1145/2858036.2858498.

Card, S. K., Newell, A., & Moran, T. P. (1983). *The psychology of human-computer interaction*. L. Erlbaum Associates Inc..

Cheng, Y., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning, 49*(3), 417–446. https://doi.org/10.1111/0023-8333.00095.

Connor, U. (1996). Contrastive rhetoric: Cross-cultural aspects of second-language writing. Cambridge University Press,.

Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing, 18*(2), 119–135. https://doi.org/10.1016/j.jslw.2009.02.002.

Cumming, A. (2009). Assessing academic writing in foreign and second languages. *Language Teaching, 42*(1), 95–107. https://doi.org/10.1017/S0261444808005430.

Curran, M., Draus, P., Maruschock, G., & Maier, T. (2013). Student perceptions of project essay grade (PEG) software. *Issues in Information Systems, 14*, 89–98.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction, 22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001.

Demmans Epp, C. (2016). English language learner experiences of formal and informal learning environments. In*Proceedings of the sixth international conference on Learning Analytics & Knowledge - LAK '16* (pp. 231–235). https://doi.org/10.1145/2883851.2883896.

Demmans Epp, C. (2017). Migrants and Mobile technology use: Gaps in the support provided by current tools. *Journal of Interactive Media in Education, 2017*(1). https://doi.org/10.5334/jime.432.

Demmans Epp, C. (2018). Developing an adaptive Mobile tool to scaffold the communication and vocabulary Acquisition of Language Learners. Handbook of Mobile Teaching and Learning, 1–26. https://doi.org/10.1007/978-3-642-41981-2_92-1.

Demmans Epp, C., & Bull, S. (2015). Uncertainty representation in visualizations of learning analytics for learners: Current approaches and opportunities. *IEEE Transactions on Learning Technologies, 8*(3), 242–260. https://doi.org/10.1109/TLT.2015.2411604.

Demmans Epp, C., & McCalla, G. (2011). ProTutor: Historic open learner models for pronunciation tutoring. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Artificial Intelligence in Education (pp. 441–443)*. Springer.

Derwing, T. M., Waugh, E., & Institute for Research on Public Policy. (2012). Language skills and the social integration of Canada's adult immigrants. Institute for Research on Public Policy. http://www.deslibris.ca/ID/233115

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification." *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040.

DiGiovanni, E., & Nagaswami, G. (2001). Online peer review: An alternative to face-to-face? *ELT Journal, 55*(3), 263–272. https://doi.org/10.1093/elt/55.3.263.

Dikli, S., & Bleyle, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback? *Assessing Writing, 22*, 1–17. https://doi.org/10.1016/j.asw.2014.03.006.

Dixon, L. Q., Zhao, J., Shin, J.-Y., Wu, S., Su, J.-H., Burgess-Brigham, R., Gezer, M. U., & Snow, C. (2012). What we know about second language acquisition: A synthesis from four perspectives. *Review of Educational Research, 82*(1), 5–60. https://doi.org/10.3102/0034654311433587.

Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education, 63*, 380–392. https://doi.org/10.1016/j.compedu.2012.12.020.

Elouazizi, N., Birol, G., Jandciu, E., Öberg, G., Welsh, A., Han, A., & Campbell, A. (2017). Automated Analysis of Aspects of Written Argumentation. Proceedings of the Seventh International Learning Analytics & Knowledge Conference, 606–607. https://doi.org/10.1145/3027385.302948

Etikan, I. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1. https://doi.org/10.11648/j.ajtas.20160501.11.

Fang, Y. (2010). Perceptions of the computer-assisted writing program among EFL college learners. *Journal of Educational Technology & Society, 13*(3), 246–256.

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition, 32*(2), 181–201. https://doi.org/10.1017/S0272263109990490.

Gibson, A., Aitken, A., Sándor, Á., Buckingham Shum, S., Tsingos-Lucas, C., & Knight, S. (2017). Reflective writing analytics for actionable feedback. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 153–162. https://doi.org/10.1145/3027385.3027436.

Gould, J. D., Boies, S. J., & Lewis, C. (1991). Making usable, useful, productivity-enhancing computer applications. *Communications of the ACM, 34*(1), 74–85. https://doi.org/10.1145/99977.99993.

Government of Canada. (2011, April 1). *Evaluation of the Language Instruction for Newcomers to Canada (LINC) Program* [Assessments]. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/reports-statistics/evaluations/language-instruction-newcomers-canada-2010/intro.html#a2

Greene, J. A., Muis, K. R., & Pieschl, S. (2010). The role of epistemic beliefs in students' self-regulated learning with computer-based learning environments: Conceptual and methodological issues. *Educational Psychologist, 45*(4), 245–257. https://doi.org/10.1080/00461520.2010.515932.

Greer, J., McCalla, G., Collins, J. A., Kumar, V. S., & Meagher, P. (1998a). Supporting peer help and collaboration in distributed workplace environments. *International Journal of Artificial Intelligence in Education, 9*, 159–177.

Greer, J., McCalla, G., Cooke, J., Collins, J., Kumar, V., Bishop, A., & Vassileva, J. (1998b). The intelligent helpdesk: Supporting peer-help in a university course. In B. P. Goettl, H. M. Halff, C. L. Redfield, & V. J. Shute (Eds.), *Intelligent Tutoring Systems (Vol. 1452, pp. 494–503)*. Springer. https://doi.org/10.1007/3-540-68716-5_55.

Greer, J., McCalla, G., Vassileva, J., Deters, R., Bull, S., & Kettel, L. (2001). Lessons learned in deploying a multi-agent learning support system: The I-help experience. Proceedings of the international conference on artificial intelligence in education, 410–421. https://www.researchgate.net/publication/242370163_Lessons_Learned_in_Deploying_a_Multi-Agent_Learning_Support_System_The_I-Help_Experience

Guerrero, M. C. M. D., & Villamil, O. S. (2002). Activating the ZPD: Mutual scaffolding in L2 peer revision. *The Modern Language Journal, 84*(1), 51–68. https://doi.org/10.1111/0026-7902.00052.

Hajer, A., & Kaskens, A, M. (2012). *Canadian language benchmarks: english as a second language for adults* (p. 209). Citizenship and Immigration Canada. https://www.canada.ca/content/dam/ircc/migration/ircc/english/pdf/pub/language-benchmarks.pdf

IELTS. (2019). https://www.ielts.org/

Immigration, Refugees and Citizenship Canada. (2013, October 18). *ARCHIVED – Backgrounder—Language Instruction for Newcomers to Canada (LINC) Program* [backgrounders]. Gcnws. https://www.canada.ca/en/immigration-refugees-citizenship/news/archives/backgrounders-2013/language-instruction-newcomers-canada-linc-program.html

Immigration, Refugees and Citizenship Canada. (2016, March 8). Annual Report to Parliament on Immigration,2015. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/publications-manuals/annual-report-parliament-immigration-2015.html#fn

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0.

Jones, E. (2008). Predicting performance in first-semester college basic writers: Revisiting the role of self-beliefs. *Contemporary Educational Psychology, 33*(2), 209–238. https://doi.org/10.1016/j.cedpsych.2006.11.001.

Kay, J. (1999). *A Scrutable user Modelling Shell for user-adapted interaction*. Department of Computer Science: University of Sydney.

Kerly, A., Hall, P., & Bull, S. (2007). Bringing chatbots into education: Towards natural language negotiation of open learner models. *Knowledge-Based Systems, 20*(2), 177–185. https://doi.org/10.1016/j.knosys.2006.11.014.

Koul, R., Roy, L., Kaewkuekool, S., & Ploisawaschai, S. (2009). Multiple goal orientations and foreign language anxiety. *System, 37*(4), 676–688. https://doi.org/10.1016/j.system.2009.09.011.

Krashen, S. D. (1983). Principles and practice in second language acquisition. *TESOL Quarterly, 17*(2), 300–305. https://doi.org/10.2307/3586656.

Lally, C. G. (2000). First language influences in second language composition: The effect of pre-writing. *Foreign Language Annals, 33*(4), 428–432. https://doi.org/10.1111/j.1944-9720.2000.tb00623.x.

Liaqat, A., Axtell, B., Munteanu, C., & Demmans Epp, C. (2018). *Contextual inquiry, participatory design, and learning analytics: An example*, *Companion proceedings of the 8th international conference on Learning Analytics & Knowledge*. Sydney: Learning Analytics & Knowledge Conference.

Liou, H.-C., & Peng, Z.-Y. (2009). Training effects on computer-mediated peer review. *System, 37*(3), 514–525. https://doi.org/10.1016/j.system.2009.01.005.

Mabbott, A., & Bull, S. (2006). Student preferences for editing, persuading, and negotiating the open learner model. In M. Ikeda, K. D. Ashley, & T.-W. Chan (Eds.), *Intelligent Tutoring Systems (Vol. 4053, pp. 481–490)*. Springer. https://doi.org/10.1007/11774303_48.

Mathews-Aydinli, J. (2008). Overlooked and understudied? A survey of current trends in research on adult English language learners. *Adult Education Quarterly, 58*(3), 198–213. https://doi.org/10.1177/0741713608314089.

Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System, 32*(1), 21–36. https://doi.org/10.1016/j.system.2003.08.002.

Mendonça, C. O., & Johnson, K. E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly, 28*(4), 745–769. https://doi.org/10.2307/3587558.

Midgley, C., & Urdan, T. (1995). Predictors of middle school students' use of self-handicapping strategies. *The Journal of Early Adolescence, 15*(4), 389–411. https://doi.org/10.1177/0272431695015004001.

Midgley, C., Kaplan, A., & Middleton, M. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? *Journal of Educational Psychology, 93*, 77–86. https://doi.org/10.1037//0022-0663.93.1.77.

Moran, T, P. (1996). *Design rationale: concepts, techniques, and use* (J. M. Carroll, Ed.). L. Erlbaum associates Inc.

Muis, K. (2007). The role of epistemic beliefs in self-regulated learning. *Educational Psychologist, 42*, 173–190. https://doi.org/10.1080/00461520701416306.

Muis, K. R., & Franco, G. M. (2009). Epistemic beliefs: Setting the standards for self-regulated learning. *Contemporary Educational Psychology, 34*(4), 306–318. https://doi.org/10.1016/j.cedpsych.2009.06.005.

Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review, 18*(1), 3–54. https://doi.org/10.1007/s10648-006-9003-6.

Munteanu, C., Lumsden, J., Fournier, H., Leung, R., D'Amours, D., McDonald, D., & Maitland, J. (2010). ALEX: Mobile language assistant for low-literacy adults. In*Proceedings of the 12th international conference on human computer interaction with Mobile devices and services* (pp. 427–430). https://doi.org/10.1145/1851600.1851697.

Munteanu, C., Molyneaux, H., Maitland, J., McDonald, D., Leung, R., Lumsden, J., & Fournier, H. (2012). Tale of two studies: Challenges in field research with low-literacy adult learners in a developed country. *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, 489–504. https://doi.org/10.1145/2212776.2212825.

Munteanu, C., Molyneaux, H., Maitland, J., McDonald, D., Leung, R., Fournier, H., & Lumsden, J. (2014). Hidden in plain sight: Low-literacy adults in a developed country overcoming social and educational challenges through mobile learning support tools. *Personal and Ubiquitous Computing, 18*(6), 1455–1469. https://doi.org/10.1007/s00779-013-0748-x.

Murau, A. M. (1993). Shared writing: Students' perceptions and attitudes of peer review. *Working Papers in Educational Linguistics, 9*(2), 71–79.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science, 37*(4), 375–401. https://doi.org/10.1007/s11251-008-9053-x.

Orji, F, A., Vassileva, J., & Greer, J. (2018). Personalized persuasion for promoting students' engagement and learning. Proceedings of the third international workshop on personalization in persuasive technology co-located with the 13th international conference on persuasive technology, PPT@PERSUASIVE 2018, Waterloo, Canada, April 16, 2018., 77–87. http://ceur-ws.org/Vol-2089/8_Orji.pdf

Orji, F, A., Oyibo, K., Orji, R., Greer, J., & Vassileva, J. (2019). Personalization of persuasive Technology in Higher Education. *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization - UMAP '19*, 336–340. https://doi.org/10.1145/3320435.3320478.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and Bias in human use of automation: An Attentional integration. *Human Factors, 52*(3), 381–410. https://doi.org/10.1177/0018720810376055.

Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing, 8*(3), 265–289. https://doi.org/10.1016/S1060-3743(99)80117-9.

Peerceptiv—Data Driven Peer Assessment. (n.d.). Panther Learning. Retrieved January 21, 2018, from http://www.peerceptiv.com/wordpress/

Pekrun, R., Vogl, E., Muis, K., & Sinatra, G. (2017). Measuring emotions during epistemic activities: The Epistemically-related emotion scales. *Cognition and Emotion, 31*(6), 1268–1276. https://doi.org/10.1080/02699931.2016.1204989.

Perin, D., & Lauterbach, M. (2018). Assessing text-based writing of low-skilled college students. *International Journal of Artificial Intelligence in Education, 28*(1), 56–78. https://doi.org/10.1007/s40593-016-0122-z.

Phillips, J. M., & Gully, S. M. (1997). Role of goal orientation, ability, need for achievement, and locus of control in the self-efficacy and goal—Setting process. *Journal of Applied Psychology, 82*(5), 792–802. https://doi.org/10.1037/0021-9010.82.5.792.

Phirangee, K., Demmans Epp, C., & Hewitt, J. (2016). Exploring the relationships between facilitation methods, students' sense of community and their online Behaviours. *Online Learning, 20*(2). https://doi.org/10.24059/olj.v20i2.775.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*(3), 544–555.

Pintrich, P, R., Smith, D, A, F., Garcia, T., & McKeachie, W, J. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)* (guides - non-classroom use (055) – tests/evaluation instruments (160) NCRIPTAL-91-B-004; p. 76). National Center for research to improve postsecondary teaching and learning. https://eric.ed.gov/?id=ED338122

Pintrich, P. R., Conley, A. M., & Kempler, T. M. (2003). Current issues in achievement goal theory and research. *International Journal of Educational Research, 39*(4), 319–337. https://doi.org/10.1016/j.ijer.2004.06.002.

Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge University Press.

Rohman, D. G. (1965). Pre-writing the stage of discovery in the writing process. *College Composition and Communication, 16*(2), 106–112. https://doi.org/10.2307/354885.

Roscoe, R. D., & McNamara, D. S. (2013). Writing pal: Feasibility of an intelligent writing strategy tutor in the high school classroom. *Journal of Educational Psychology, 105*(4), 1010–1025. https://doi.org/10.1037/a0032340.

Roscoe, R. D., Wilson, J., Johnson, A. C., & Mayra, C. R. (2017). Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation. *Computers in Human Behavior, 70*, 207–221. https://doi.org/10.1016/j.chb.2016.12.076.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*, 498–504. https://doi.org/10.1037/0022-0663.82.3.498.

Schommer, M. (1998). The influence of age and education on epistemological beliefs. *British Journal of Educational Psychology, 68*(4), 551–562.

Settings IELTS entry scores. (n.d.). Retrieved May 29, 2018, from https://www.ielts.org/ielts-for-organisations/setting-ielts-entry-scores

Shafaei, A., & Nejati, M. (2008). *Global practices of language teaching: Proceedings of the 2008 international online language conference (IOLC 2008)*. Universal-Publishers.

Shahrour, G., & Bull, S. (2008). Does 'notice' prompt noticing? Raising awareness in language learning with an open learner model. In W. Nejdl, J. Kay, P. Pu, & E. Herder (Eds.), *Adaptive Hypermedia and Adaptive Web-Based Systems (Vol. 5149, pp. 173–182)*. Springer. https://doi.org/10.1007/978-3-540-70987-9_20.

Shao, X. (2015). On written corrective feedback in L2 writing. *English Language Teaching, 8*(3). https://doi.org/10.5539/elt.v8n3p155.

Shum, S. B., Sándor, Á., Goldsmith, R., Wang, X., Bass, R., & McWilliams, M. (2016). Reflecting on reflective writing analytics: Assessment challenges and iterative evaluation of a prototype tool. In*Proceedings of the sixth international conference on Learning Analytics & Knowledge* (pp. 213–222). https://doi.org/10.1145/2883851.2883955.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly, 27*(4), 657–677. https://doi.org/10.2307/3587400.

Soegaard, M., & Friis Dam, R. (2013). The encyclopedia of human-computer interaction, 2nd Ed.

Smith, J. A., Harre, R., & Langenhove, L. V. (1995). Rethinking Methods in Psychology. SAGE

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14*(3), 153–173. https://doi.org/10.1016/j.jslw.2005.05.002.

Storch, N. (2011). Collaborative writing in L2 contexts: Processes, outcomes, and future directions. *Annual Review of Applied Linguistics, 31*, 275–288. https://doi.org/10.1017/S0267190511000079.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics*. Oxford University Press.

Tang, G. M. (1999). Peer response in ESL writing. *TESL Canada Journal, 16*(2), 20. https://doi.org/10.18806/tesl.v16i2.716.

Taylor, R. (2012). *Review of the Motivated Strategies for Learning Questionnaire (MSLQ) Using Reliability Generalization Techniques to Assess Scale Reliability* [Education Foundation, Leadership, and Technology]. https://etd.auburn.edu/handle/10415/3114

Tsourounis, S., & Demmans Epp, C. (2016). Learning dashboards and Gamification in MALL: Design guidelines in practice. *The International Handbook of Mobile-Assisted Language Learning*, 370–398.

Ullmann, T, D. (2017). Reflective writing analytics—Empirically determined keywords of written reflection. *LAK '17 Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 163–167. http://oro.open.ac.uk/48840/

Vassileva, J. (2008). Toward social learning environments. *IEEE Transactions on Learning Technologies, 1*(4), 199–214. https://doi.org/10.1109/TLT.2009.4.

Vassileva, J., Greer, J., McCalla, G., Deters, R., Zapata, D., Mudgal, C., & Grant, S. (1999). A multi-agent design of a peer-help environment. In*Proceedings of the international conference on artificial intelligence in education* (pp. 38–45).

Vassileva, J., McCalla, G., & Greer, J. (2016). From small seeds grow fruitful trees: How the PHelpS peer help system stimulated a diverse and innovative research agenda over 15 years. *International Journal of Artificial Intelligence in Education, 26*(1), 431–447. https://doi.org/10.1007/s40593-015-0073-9.

Vines, J., Blythe, M., Dunphy, P., Vlachokyriakos, V., Teece, I., Monk, A., & Olivier, P. (2012). Cheque mates: Participatory Design of Digital Payments with eighty Somethings. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1189–1198. https://doi.org/10.1145/2207676.2208569.

Vines, J., Clarke, R., Wright, P., McCarthy, J., & Olivier, P. (2013). Configuring participation: On how we involve people in design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, 429. https://doi.org/10.1145/2470654.2470716.

Wagner, R, K., Muse, A, E., & Tannenbaum, K, R. (Eds.). (2007). *Vocabulary acquisition: Implications for reading comprehension.* (pp. xvi, 301). Guilford Press.

Waller, L., & Papi, M. (2017). Motivation and feedback: How implicit theories of intelligence predict L2 writers' motivation and feedback orientation. *Journal of Second Language Writing, 35*, 54–65. https://doi.org/10.1016/j.jslw.2017.01.004.

Wanderley, L., & Demmans Epp, C. (2020). Identifying negative language transfer in writing to increase English as a second language learners' metalinguistic awareness. Bringing Together Writing Tool Design, Writing Analytics and Writing Pedagogy Workshop at the International Conference on Learning Analytics and Knowledge.

Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing, 15*(3), 179–200. https://doi.org/10.1016/j.jslw.2006.09.004.

Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching, 49*(4), 461–493. https://doi.org/10.1017/S0261444816000161.

Zapata-Rivera, J, D., & Greer, J. (2002). Construction and inspection of learner models. *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community*, 495–496. http://dl.acm.org/citation.cfm?id=1658616.1658687

Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of Second Language Writing, 4*(3), 209–222. https://doi.org/10.1016/1060-3743(95)90010-1.

Zhang, H., Magooda, A., Litman, D., Correnti, R., Wang, E., Matsmura, L. C., Howe, E., & Quintana, R. (2019). eRevise: Using natural language processing to provide formative feedback on text evidence usage in student writing. *Proceedings of the AAAI Conference on Artificial Intelligence, 33*, 9619–9625. https://doi.org/10.1609/aaai.v33i01.33019619.