Investigating Assessment as Learning in Second Language Writing: A Qualitative Research Perspective

Ricky Lam

Abstract
Using assessment to improve learning has become a global phenomenon. Research has shown that applying assessment for learning (AfL) in classroom contexts could increase learning gains. Despite this, not much is done to substantiate how assessment as learning (AaL), a subset of AfL, has developed to support English-as-a-second-language (ESL) pupils’ learning of writing because teacher-centered pedagogy remains prevalent in the ESL educational landscape. In this article, I first introduce the idea of AaL. Second, I review five AaL studies with a focus on what research methods they adopted and evaluate whether these methods were applied appropriately. Third, I suggest four qualitative research methods that are suitable to track how AaL could facilitate pupils’ continued writing development. Fourth, I present a case study, exemplifying how I use three research methods to investigate young adolescents’ (Grade 7) mastery of AaL skills in a forthcoming project. This article ends with methodological implications, which enlighten scholars and teachers to investigate AaL in second language writing from a qualitative research perspective.

Keywords
qualitative research methods, assessment for learning, assessment as learning, L2 writing

Introduction
In education, scholars generally adopt three common methods to conduct empirical studies, namely quantitative, qualitative, and mixed-methods approaches (Creswell & Creswell, 2017). While these three research approaches are commonplace, this article only focuses on the qualitative research approach with the following rationale. First, a majority of studies researching assessment for/as learning (AfL/AaL) has adopted the quantitative approach, chiefly measuring pupils’ achievement outcomes. Because of this, we have limited conceptual understanding of pupils’ learning processes when they engage in AaL practices. Second, the importance of qualitative methods appears to be downplayed in second language (L2) writing research, particularly of AaL studies (Lee, 2017). Third, to explore pupils’ developmental trajectories in writing, researchers may find the qualitative approach epistemologically and methodologically convincing, given that they inquire into the process of learning rather than the product of learning. The overall aim of this article is to identify appropriate qualitative research methods that can be used to investigate AaL in L2 writing classrooms with theoretical rationale, research evidence, and a case study of my forthcoming project.

This article is divided into five parts. First, I provide the definitions of assessment of learning (AoL), AfL, and AaL in the L2 contexts. I further discuss the relationship between AfL and AaL and the role of AaL in L2 writing. Second, I review five AaL studies with a focus on what research methods they adopted and evaluate whether these methods investigated AaL in a dependable and trustworthy manner. Third, based upon the reflexive and longitudinal nature of AaL, I propose four qualitative research methods that are considered innovative to examine and understand the pedagogical impact of AaL on pupil learning, especially in ethnographic and longitudinal studies.

Fourth, I report on my forthcoming AaL project to rationalize why I adopt three qualitative research methods as mentioned in the previous section and to analyze the pros and cons of these methods when put into practice. Lastly, this article
implies that it is significant to identify *ecologically valid* research tools to investigate the middle-term and long-term impact of AaL on pupil learning. In this article, I intend to deepen writing teachers’ and scholars’ conceptual understanding and knowledge, so that they can investigate AaL with proper methods. Although I advocate qualitative methods in this article, I have no intention to downplay the role of quantitative or mixed-methods approaches in AaL scholarship, particularly when researchers measure the achievement outcomes of AaL-induced pedagogy. The next section is about the definition of AaL and its role in L2 writing.

**AaL and AfL in L2 Writing**

In assessment literature, there are three purposes of assessment, namely AoL, AfL, and AaL. AoL refers to using assessment to judge pupil learning near the end of an instructional unit. AfL is a learning-oriented process, facilitating pupils to receive, generate, and act upon feedback to improve learning. This process, planned or unplanned, chiefly takes place in the classroom, triggered by a range of contingent pedagogical moments, which in turn inform effective teaching and learning (Bennett, 2011; Wiliam, 2011). Likewise, AaL capitalizes on pupil ability to self-assess, self-monitor, and self-regulate with reference to the pupils’ set goals in order to close their learning gaps (Bennett & Gitomer, 2009; Clark, 2012). Regardless of their divergent nature, these three purposes of assessment are not mutually exclusive although teachers and pupils tend to pay undue attention to AoL more than AfL/AaL. As opposed to some Anglophone societies, teachers and pupils are typically exam-driven and grade-conscious in L2 contexts, provided that exit exams are considered a major gateway to upward mobility (Carless & Lam, 2014). Scholars have reported that within L2 classrooms and some developing countries, the principles of AfL and AaL are seldom practiced since they are encouraged but not mandated by the government and commonly perceived as a barrier to getting high scores (Lam, 2018a). In these contexts, AfL/AaL appears in policy and curriculum documents, but training for teachers is lacking. In fact, with effective planning, AfL and AaL could be seamlessly integrated into AoL to help pupils improve exit exam results (Sadeghi & Rahmati, 2017). While Earl (2013) has pointed out that AaL is a subset of AfL, other scholars believe that AaL could be considered a stand-alone instructional approach, which transforms pupils into committed, critical, and self-regulated learners who create new learning by synthesizing prior and newly learnt knowledge (Gottlieb, 2016).

Notwithstanding the differences between AfL and AaL, they have one feature in common, which is the use of feedback, be it internal or external, to develop pupil metacognitive knowledge and awareness to perform regulation of learning (Andrade & Brookhart, 2016; Earl & Katz, 2008). This commonality on feedback enactment makes AfL and AaL inseparable. That said, in AfL, pupils generally count on multiple sources of feedback, for example, self-, peer-, teacher, or electronic feedback to make adjustments to their learning, whereas in AaL, pupils chiefly mobilize internally generated feedback to plan, monitor, evaluate, and revise their learning independently or at times collaboratively with teachers (Wiliam & Thompson, 2007). In brief, AfL is public, interactive, and pedagogical, highlighting the essence of community of practice. AaL is relatively private, individualistic, and reflexive, capturing the key feature of self-regulated learning (Allal, 2019; Lam, 2018b).

Similar to L1 writing, the role of AaL in L2 writing chiefly focuses on two aspects, namely metacognition and formative assessment processes. Metacognition is concerned with equipping pupils’ knowledge and skills in self-assessing, self-monitoring, and self-regulating their learning by addressing the three questions: Where am I going? How am I going? What is the next step of learning? (Perrenoud, 1998). Major components of metacognition could be flexibly incorporated into pre-writing, while-writing, and postwriting stages because these procedures neatly dovetail with recursive composing processes including planning, drafting, self-reflecting, and revising (four AaL strategies in the writing classroom, cf. Lee, 2017). As pointed out by Lee (2016), AaL is to put pupil learning at the center of classroom writing assessments. One prominent example of this pupil-centered AaL is writing portfolios, in which pupils are encouraged to set goals with rubrics, self-monitor writing development with these rubrics, self-assess writing progress with feedback, and revise works in progress whenever necessary. The benefits of developing pupil metacognitive awareness and capacity for AaL include enhanced motivation, ownership, and autonomy in learning writing (Mak & Wong, 2018). Nevertheless, there are possible pedagogical issues of applying metacognition in L2 settings since teachers and pupils may not have sufficient exposure to self-reflective practices nor were they used to a process-oriented writing approach. After all, the product-oriented approach to teaching and learning writing remains predominant owing to limited teacher assessment training and a larger exam culture (Lee, 2017).

The second aspect of AaL in L2 writing is formative assessment processes, in which teachers take the lead role in clarifying learning intentions and sharing success criteria, designing tasks which elicit evidence of learning, interpreting the evidence in relation to the success criteria, and providing learners with descriptive feedback which extends and deepens pupil learning (Wiliam, 2011). The formative assessment processes within AaL provide pupils with timely input to self-regulate learning after they identify the gap between their desired and existing levels of performances in writing (Heritage, 2010). What metacognition and formative assessment processes imply in L2 writing settings is that teachers are required to play a lesser role, whereas pupils are expected to take center stage in self-regulating their learning, which may pose a challenge to pupils being used to teacher-centered, textbook-bound, and didactic instructional approaches. In reality, teachers may find it hard to teach less and let pupils learn more reflectively through AaL-focused tasks. To complicate this pedagogical matter further, explicit instruction in metacognitive strategies is proved to be few and far between and professionally taxing,
if not entirely impossible (Nguyen & Gu, 2013). Hence, AaL in L2 classrooms perhaps calls for a teacher- or technology-assisted version because L2 pupils may need more support to uptake those metacognitive skills. Besides providing a sound understanding of what AaL entails, I demonstrate how other scholars investigated these two AaL aspects—metacognition and formative assessment processes—within an L2 writing context by reviewing pertinent research studies. Since AaL is reflexive, idiosyncratic, and internally driven, using conventional methods like elicitation may not prove to be a wise option. The subsequent section critically reviews the research methods adopted in five AaL studies.

**Review of AaL Studies and Their Research Methods**

The selection of these studies began with Lee’s (2017) theoretical discussion on the role of AaL in L2 classroom contexts (cf. Chapter 4 of her monograph). In her discussion, she has included eight empirical studies. Among these, five of the eight studies have fulfilled the following three preset selection criteria: (1) studies with empirical data and a succinct methodology section; (2) studies with inclusion of AaL components such as self-assessment, self-reflection, and self-monitoring; and (3) studies utilizing L2 writing as a research context. Upon completion of the selection, I reread the five studies twice and categorized the attributes of AaL to match with those defined in Earl’s (2013) book, which emphasized students being critical thinkers, critical connectors, and knowledge contributors to align assessment and learning. This verification procedure enhances the trustworthiness of the review and the argument in favor of qualitative research methods in the AaL scholarship.

The chosen studies represent the adoption of a range of methodologically appropriate research strategies to investigate AaL in L2/English-as-a-foreign-language (EFL) writing. However, it is almost impossible to include more AaL studies due to space limit. Although the reviewed studies are mostly in higher education, the focus of this article is to evaluate the suitability of their research methods not the level of schooling and find out whether these methods are applicable to my case study examining how young adolescents develop and retain their reflective capacities. With that said, I remain vigilant and keep on scrutinizing the transferability of these methods for learners at different developmental stages. The first reviewed article is Feltham and Sharen’s (2015) research project, which examined whether 17 participants changed their mindsets after experiencing training in text revision and the extent to which they benefited from the classroom interventions. The study took place in a second-year undergraduate course in one Canadian university. In the study, the participants received 6 hr of training in library research, critical thinking, writing processes, feedback literacy, and revision strategies. Two research methods were employed, including (a) a pretest and a posttest about learning goals, metacognitive strategy use, behavioral strategies, self-regulation strategies, and writing self-efficacy (from three tried-and-tested instruments) and (b) text analysis (grading of each participant’s draft and final versions of the assignment). I concur that utilizing pre- and posttests to measure the “fixed” and “growth” mindsets is methodologically legitimate. However, the application of self-regulated strategies in assignment revision may not be accurately reflected because only one interim draft rather than multiple rounds of revision was required from each participant. Further, the authors only measured the writing products (average scores of the two drafts) without looking into what metacognitive composing strategies the participants employed when revising. For instance, introspective verbal reports could be considered to elicit the participants’ metacognitive composing behaviors.

The second paper is Xiang’s (2004) quasi-experimental study, which is about using self-monitoring in undergraduate students’ English writing. The study investigated whether the participants could be trained to self-monitor skillfully, the impact of self-monitoring on writing, and the participants’ attitudes toward self-monitoring. Two classes of English majors were divided into an experimental group with self-monitoring training and a control group. In the experimental group, two training sessions were provided to coach the participants to critique their essays in terms of content, organization, and form via annotations. The author adopted four research methods including a pretest and a posttest on writing, four writing tasks, a questionnaire, and interviews. Like the previous study, the author chiefly evaluated the outcomes of self-monitoring by statistically comparing the participants’ annotated texts, essay scores, and learning gains in both experimental and control groups. Nonetheless, the processes of self-monitoring were not fully researched, such as what decisions the participants made while annotating drafts, when they performed annotations, at home or in class, and what support they needed for self-monitoring, namely self-, peer-, teacher feedback, or other source texts. The author should require the participants to keep a journal or a narrative frame, which tracks their self-monitoring experiences in greater detail.

Huang’s (2015) study, our third reviewed paper, examined the effect of goal-setting on students’ revision of writing in an EFL context. The study took place in a Taiwanese university with 111 participants. The paper inquired whether student writing improved from each draft to its revision in the control, goal, and goal+ groups and whether the quality of the three drafts or three revisions improved in a writing course. The control group involved no goal-setting. The goal group referred to goals set by learners before revision, and the goal+ group involved learners to set goals with revision strategies provided to them. The author adopted text analysis to determine the quality of three drafts and three revisions. Although rigorous statistical methods were used, including Pearson correlation, an analysis of variance, and an analysis of covariance, these quantitative methods could verify the impact of goal-setting on students’ writing linguistically. Hence, text analysis might not enlighten readers on how the participants reacted to affective and cognitive aspects of revision. At present, we know little about how goal-setting affected students’ revision behaviors. Nor do we understand what revision strategies students were able to
master after they set these learning goals. Alongside text analysis, the authors may consider think-aloud protocols (Wong, 2005) or stimulated recalls (Lee, 2020) to identify specific revision strategies adopted by the participants.

The fourth paper I reviewed is Hawe and Dixon’s (2014) study about how three school teachers developed Grades 5, 7, and 8 pupils’ evaluative knowledge and productive skills when teaching writing. The project is the second phase of a larger project, which examined New Zealand teachers’ beliefs and knowledge about feedback and to investigate their perceptions of practice. In Phase 2, three teacher participants were purposefully selected because they believed that pupils should take an active role in the feedback process by practicing peer appraisals, self-assessment, or goal-setting in relation to success criteria. Qualitative data collection involved observations (one teaching unit of writing for each participant), postobservation interviews, and documents (e.g., language unit plans, worksheets, and success criteria). While the study has meticulously demonstrated how the three teachers introduced AaL-related pedagogical strategies to develop pupils’ evaluative and productive expertise in writing, the authors may invite the teacher participants to keep a narrative frame of their espoused beliefs about learner-centered feedback processes or probably analyze parts of pupils’ drafts and final writings to see whether there is any uptake of metacognitive capacity after explicit AaL instruction. After all, the inner workings of the teachers’ and pupils’ minds about the application of AaL remain unexplored in the study.

The last paper I reviewed is Davies et al.’s (2011) project, which described how 27 tertiary-level teachers engaged in a wiki-writing assignment emphasizing AaL. The study aimed to investigate how the wiki-writing program could encourage the participants to share their assessment knowledge, philosophies, and practices via an online learning community in an advanced diploma course. Prior to data collection, the authors provided the participants with 1-day workshop to familiarize themselves with what would be expected from the study, including contents of writing, interaction modes, ways to support one another, and so on. Four research methods were deployed in the study, namely the first and second authors’ reflection, observation by the third author, the participants’ wiki-writing, and a postprogram questionnaire. Although the participants’ perceptions about wiki-writing were elicited, we had no idea of how they collaborated, became more self-aware and metacognitive in real time. Further, for learning inquiry, it appears that the study just summarized the participants’ self-reported data (in four subpoints) without identifying the “how” aspect of AaL, say moments or episodes that witnessed self-assessment, self-monitoring, and self-regulatory behaviors. In fact, the participants’ reflexivity, interaction patterns with (coursemates), and metacognitive acts could be captured online or off-line via document analysis.

Taken together, the five reviewed studies are relatable to the main goal of this article, which is to identify appropriate qualitative methods to investigate AaL in the L2 writing setting. Nonetheless, the reviewed studies seem to unanimously focus on the effects of AaL on participants’ learning outcomes—the product of learning (i.e., scores and means) and utilized the participants’ self-reported data as major findings. While these emphases were legitimate, more could be done to understand the process-oriented nature of AaL in a longitudinal manner, namely self-monitoring patterns, self-regulatory behaviors, or metacognitive acts by way of other qualitative methods like narrative frames, introspective/retrospective verbal reports, journals, and document analysis of multimodal texts, for example, online discussion forums or digital peer feedback (Barkhuizen & Wette, 2008; Bowen, 2009; Lee, 2017). Despite these alternatives, the use of quasi-experimental or qualitative designs to examine AaL is methodologically valid and conceptually significant (e.g., Butler & Lee, 2010). The next section proposes four qualitative research methods, which may help scholars explore AaL in L2 writing more readily.

**Qualitative Research Methods for AaL**

Based on the above review, there is a commonality between these studies. The authors primarily adopted a quantitative achievement-driven approach, which only considered the outcome and not the process. Because of this, I suggest that AaL studies be best investigated by qualitative research methods because they can capture the unique characteristics and processes of AaL with high fidelity. The four qualitative research methods for AaL include (1) narrative frames, (2) retrospective verbal reports, aka verbal protocols, (3) the draw-and-talk technique, and (4) social media tools, namely Facebook. There are three justifications for introducing these methods. First, they are “tried and tested” and considered “innovative” in leading qualitative research studies (cf. Wiles et al., 2011). Second, pupil application of metacognitive composing strategies (e.g., self-monitoring) in authentic writing tasks like how and why can be investigated by the Methods (2) and (3), as these methods can examine beliefs, behaviors, and emotional experiences. Third, these alternative qualitative methods are less geographically restrictive in the sense that researchers can collect data off campus or simply online as in the Methods (1) and (4).

**Narrative Frames**

Narrative frames stem from a methodological approach called narrative inquiry, in which researchers examine informants’ narratives to (co)construct realities (Barkhuizen & Wette, 2008). Narrative inquiry is broadly defined as a way of conducting research, which emphasizes using a person’s past, present, or future stories as evidence via recounting, reflecting, and negotiating about their experiences of life (Barkhuizen et al., 2014). Since the turn of the century, narrative inquiry has become a thriving qualitative methodology used by postmodernist scholars, who prefer researching people’s agency, identities, and experiences to validating preestablished hypotheses by sophisticated statistical tools (Benson, 2013). Narrative inquiry theoretically hinges upon a notion of narrative knowledging, implying that this methodology is more than a set of
techniques to collect data for understanding a person’s lives. Instead, it generates knowledge and develops grounded theories as a unified whole because the research processes are dynamic, evolving, and meaning-making (Barkhuizen, 2011).

One prominent technique of narrative inquiry is narrative frames, which refer to “written story templates consisting of incomplete sentences or blank spaces of varying lengths” (Barkhuizen, 2015, p. 178) for informants to fill out their own vignettes of lived experiences. These templates are usually designed in skeletal form. When completed, they provide researchers with a full version of the informants’ narratives and reflections. Narrative frames could be simultaneously adopted as a research method and a learning-oriented task when investigating AaL since its very nature is metacognitive and introspective, facilitating the informants to recapitulate how they self-regulate the learning of writing. Although narrative frames are exploratory, flexible, and informative, they have certain setbacks that may prevent researchers from collecting dependable data. For instance, researchers may include too many topics in one frame, which makes informants unable to retell their stories precisely. The format of the frame may be linguistically incompatible with the respondent’s cultural writing conventions (cf. Barkhuizen, 2014). Scholars also need to pay close attention to research ethics, given that informants may not know that their stories will reach a wider audience. Despite these limitations, narrative frames remain valid and pedagogical, especially when used as parts of AaL instructional materials (e.g., postwriting reflection pieces).

Retrospective Verbal Reports

Verbal reports refer to a research method, in which participants report back to the researcher orally when they engage in a cognitive or a linguistic task (Punch, 2014). There are two types of verbal reporting, including “think-aloud” and stimulated recalls (Creswell & Creswell, 2017). The former requires participants to verbalize everything on their minds while completing a task, namely their thoughts and actions during the revision of a draft. The latter demands participants to recollect their mental processes and learning strategies right after they complete a writing task. For example, participants were asked how and why they rephrased certain sentences through a videotaped segment of their real-time composing behaviors (Hyland, 2015). In Lam’s (2015) study, he used stimulated recalls to find out which self-regulated writing strategies the four informants employed when they drafted an argumentative essay. Similarly, Lee (2020) adopted stimulated recalls to identify 48 cognitive strategies generated by six Grade 9 students, who revised their compositions in response to an automated content feedback system.

Think-aloud and stimulated recalls are appropriate for researching AaL in L2 writing. It is because the procedures of self-monitoring and self-regulating in AaL call for a method, which truthfully reflects actual mental processes in action such as planning, reorganizing, and problem-solving. This metacognitive knowledge may not be easily identified by elicitation (interviews and questionnaires) and text samples (document analysis). Further, introspective and retrospective protocols are considered more reliable than self-reported data since participants directly tell the researcher how and why they engage in a writing task almost immediately (Hyland, 2016). Notwithstanding these benefits, verbal reports have limitations. In Hu and Gao’s (2017) study, they stated that there were two methodological drawbacks about think-aloud. First, verbal reports created reactivity effects because reporting the cognitive/metacognitive processes (as in reading) would alter the processes and outcomes being reported owing to language choice and age (Greene & Azevedo, 2009). Second, verbal reports might have the issue of veridicality since the cognitive/metacognitive processes were not accurately and completely described due to task complexity. According to Hu and Gao’s (2017) recommendations, provision of appropriate training (in performing cognitive tasks) and prompting (by scaffolded questions) would mitigate the reactivity effects and maximize veridicality respectively.

Draw-and-Talk Technique

The draw-and-talk technique is considered a child-friendly research method, which collects young children’s (aged below 5) views of their learning experiences. A researcher observes a child who is drawing and talks to her concurrently about a topic under investigation. This method derives from a notion that under adult guidance, children coconstruct meanings of their lived experiences through a blend of pictorial and conversational data (Cox, 2005). The theoretical basis of this method is that during the drawing process, children externalize their inner thinking and make sense of their perceptions via visual images (Holliday et al., 2009). The technique also reinforces the role of coconstruction in the learning process, where children are provided with the right level of scaffolding to create new meanings of their learning, understanding, and experiences with adults (Bruner, 1983). Carless and Lam (2014) investigated the first-to third-graders’ conceptions of English assessment in Hong Kong by interviews and the draw-a-picture technique. Similarly, Harris et al. (2009) adopted the same technique to identify student perceptions of assessment in New Zealand schools.

Applied in L2 writing research, the draw-and-talk technique is ideal for identifying children’s cognitive and psychological aspects of AaL, which is not easily researched by conventional language-based interviews or questionnaires. This technique provides researchers with rich and in-depth insiders’ perspectives on how children set goals, monitor progress, and adjust learning through their drawing. A collection of these pictures plus synchronous conversations may develop a better understanding of why children made particular assessment decisions and adopted certain coping strategies while completing their writing tasks (e.g., sentence writing and picture compositions). Tay-Lim and Lim (2013) stated that the components of drawing and talking should go hand in hand in order to capture how children thought when they were drawing. If talking took place after drawing, children would have heavy cognitive load to
recall their memories. Tay-Lim and Lim further recommended using open-ended questions to promote deeper reflection on AaL experiences and endorsement phrases to value children’s ideas in the conversations. Despite the advantages of the technique, researchers should be aware of children’s emotional experiences and the power relationship with children.

Social Media Tools (Facebook)

Using social media for language teaching and learning has been in place for a decade. Pedagogical use of Wiki and weblogs to promote AaL in writing has been expanding (Davies et al., 2011). Owing to the popularity of social media, scholars utilize them as a tool to collect in-depth data for qualitative longitudinal research, such as Facebook (Davidson et al., 2016). According to those Facebook-based studies, researchers adopted the Timeline function as a prompt to elicit stories or “critical” moments for addressing the research problems. For instance, in Duguay’s (2014) project, she invited participants to navigate their Facebook accounts while responding to her interview questions, treating the digital evidence as a form of memory recall via photos and participants’ posts. Likewise, Gehl (2014) pointed out that researchers tend to examine participants’ cloud-based archives (e.g., Facebook Timelines) more than their print archives like physical photo albums. To this end, social media play a central role in helping scholars retrieve autobiographical data for longitudinal studies.

In Robards’s and Lincoln’s (2017) study, they investigated 34 young adults’ growing up stories in England and Australia. The authors adopted a “scrolling back method” on Facebook Timelines when interviewing the informants. In this method, the authors first asked the informants to note down their fateful moments for discussion, followed by scrolling back on their Facebook Timelines. This technique is equally viable for researching AaL owing to the following reasons. First, students can create their Facebook Timelines to record significant moments when they self-regulate their learning of writing. Second, students may consider Facebook Timelines as part of writing evidence to be included in their digital portfolios. Third, the Timeline function can help students recall memory promptly while they are self-monitoring their personal records of lived experiences in learning writing because the practice of scrolling enriches researchers’ and pupils’ understanding of sustained narrative texts. Regardless of its value, the scrolling back method may bring about ethical issues such as confidentiality and anonymity, especially when the boundary between public and private on Facebook is blurred (Zimmer, 2010). Hence, researchers should observe privacy when adopting this method. Another issue of using Facebook for AaL studies is that pupils may either mask their real personas or fabricate their metacognitive composing behaviors, which could seriously bias any conclusions drawn. The subsequent section illustrates three qualitative research methods adopted in my recent research project, which investigates how Grade 7 pupils in Hong Kong master AaL skills over the course of 2 academic years.

A Forthcoming AaL Project: Showcasing Use of Qualitative Methods

Internationally, teachers and scholars recognize that AfL can support and improve learning if applied appropriately (Black & Wiliam, 2018). Although there is no shortage of AfL studies, they are mainly about how teachers experimented with various AfL strategies in L1 classroom settings and how they handled those challenges when attempting AfL (e.g., Hawe & Dixon, 2014). Hitherto, we still know very little about how school pupils react to various AfL practices cognitively and affectively, especially in L2 environments. Likewise, when scholars advocate the value of AaL and its positive impact on learning, it appears that we lack adequate empirical evidence to understand its underlying principles, its procedures, its benefits and limitations, and its development in diverse assessment locales, namely non-Anglophone regions.

In Hong Kong, the Ministry of Education has promoted AaL for over a decade. Parallel with previous reform initiatives, the idea of using assessment to inform effective teaching and learning has extended from AfL to AaL. Based upon a new interpretation from educationalists and the Ministry of Education personnel, AaL goes beyond merely a subset of AfL. In fact, AaL is regarded as a latest and innovative pedagogical approach, which embraces the development of metacognitive thinking skills, autonomous learning, self-monitoring capacities, and metalinguistic awareness in the composing process (Curriculum Development Council [CDC], 2017). For instance, in writing, pupils are encouraged to compile a portfolio of works, which enables them to practice self-reflection and self-monitoring skills when reviewing their learning progress throughout an academic year (Lam, 2018b). That said, not much has been done to understand how Hong Kong junior secondary-level pupils conceive, master, and internalize AaL skills in the writing classroom context, where AaL remains a novelty.

Against this backdrop, my forthcoming project intends to examine how 18 seventh graders in three Hong Kong middle-range secondary schools acquire and sustain metacognitive skills for 2 school years after initial exposure to 6-week AaL-focused writing instruction. In each school, two favorably responsive, two neutrally responsive, and two less favorably responsive to AaL seventh graders will be chosen by their teachers to participate in the project. Other than metacognitive skills, I look into how the participants nurture, develop, and retain their reflective capacities over time. Data collection takes place in Semester 2 of each academic year in the project. The study follows the qualitative research paradigm, which emphasizes the use of “thick” description to report students’ lived learning experiences of AaL and its cognitive, affective, and strategic impact on the development of metacognition within an authentic L2 context. Taking these scenarios into consideration, I formulate two research questions: (1) upon 2-year exposure to AaL practices, how does AaL influence pupils’ emotion management (affect) and regulation of learning (metacognition)? And (2) to what extent do pupils retain
their reflective capacities in L2 writing contexts in the year following their training in applying AaL? In the original project, I include five research methods to answer four research questions in total, including group interviews, classroom observations, narrative frames, stimulated recalls, and Facebook Timelines. To suit the scope of this article, I discuss the last three methods in relation to their suitability to investigate AaL in context.

To address the first research question, my coinvestigator and I examine how AaL affects pupil emotion management and regulation of learning. Regarding emotion management and regulation of learning, we utilize narrative frames and stimulated recalls, respectively. When designing the narrative frames, we deliberately include prompts or blank spaces, in which the pupil informants could narrate their reflective journeys more expressively. For instance, we plan to guide students to describe how they feel about the processes of drafting the best piece of writing and the most challenging piece of writing by using emotive adjectives (e.g., proud, accomplished, frustrated) or phrases (e.g., snowed under). We also want to know how the participants manage their emotions when composing, including the use of coping strategies. In case they get stuck in drafting an argumentative essay, what metacognitive acts did they mobilize and why? How did they recount this particular self-reflection episode in their frames?

As to the regulation of learning, we will utilize stimulated recalls. To identify how students regulate the learning of writing, we originally consider using think-aloud. However, think-aloud could be cognitively exacting to young adolescents, who are unable to verbalize while thinking, composing, reflecting, and problem-solving simultaneously (Hu & Gao, 2017). The advantages of using stimulated recalls to track self-regulated learning include sufficient time to recall the composing acts, ample input to recount decision-making processes (by viewing video clips of pupils’ own writing), and adequate scaffolding to perform the recall by interacting with the researcher. In order to make the recalls of self-regulated learning dependable, the informants will perform the recalls right after completing the writing tasks within the same day. Further, we allow our informants to utter their cognitive or metacognitive strategies in L1 (i.e., Cantonese) to avoid language barriers and loss of proper expressions (Ericsson & Simon, 1993).

To answer the second research question, we explore the extent to which the informants retain reflective capacities in the year following explicit training in AaL. We use stimulated recalls and Facebook Timelines to track student retention of reflective capacities. Stimulated recalls intend to observe whether the informants carry on applying metacognitive strategies when composing, such as goal-setting, reorganizing, evaluating, and revising. The participants performed the recall tasks 3 times during Semester 2 of the second year of the project. Then, we will carefully study the transcripts of the stimulated recalls to categorize traces of self-reflection in writing. Additionally, we use Facebook Timelines to monitor whether the participants could develop reflective composing behaviors when learning writing. Scrolling back the Timelines together with the participants in an interview format, we can detect whether and why they are using AaL-related strategies through discussing specific “critical” moments. Each pupil participant is interviewed once near the end of Semester 2 of the second year. To warrant privacy, we browse the Timelines only with the informants’ consent and presence. In brief, there is always more than a method to address the same research questions. For instance, we may use the Timelines to understand student regulation of learning in the first research question and adopt narrative frames to examine the retention of reflective capacities in the second. No matter what methods we use, they should be rigorous and trustworthy in principle.

Although the aforementioned project is still in its preparation stage, we conducted a small-scale pilot study in another Grade 7 classroom besides the three participating schools in the main study. One female teacher participant tried out AaL-focused pedagogy by instructing her pupils how to plan, monitor, evaluate, and revise multiple drafts within a portfolio-based program. We attempted to answer the first research question via our pilot data. When we administered narrative frames and stimulated recalls with six pupil participants of various academic levels, we encountered some challenges including (1) their lack of vocabulary to express emotion and (2) their inability to recall critical composing behaviors. To tackle Issue no. 1, we provided two less able pupils with a list of bilingual vocabulary to assist their expressions of emotion. This scaffolded task was proved to be feasible as not every pupil had the discourse and self-confidence to share their feelings about AaL with strangers. To resolve Issue no. 2, a few pupils required us to play back the video clips with a shorter time interval. Therefore, we replayed certain segments twice or even 3 times. Whenever we observed silence, we used short questions to remind pupils of their composing acts, such as “Why did you erase ‘funny’ and then put ‘exciting’ here?” “Why didn’t you go on to explain this idea and start a new paragraph instead?” In retrospect, we should provide the participants with structured training before they filled out the narrative frames, say from short to long frames and some vocabulary items describing self-monitoring procedures. The last section discusses implications for qualitative research with a focus on AaL in L2 writing.

**Implications for Qualitative Research and Conclusion**

Considering the above discussion, I come up with four implications when researching AaL in L2 writing, including updated theoretical definition of AaL, renewed interpretation of assessment principles, use of technology in research methods, and provision of training in investigating AaL.

**Updated Theoretical Definition of AaL**

Before looking for a valid tool to research AaL, it is vital to identify a clear theoretical definition of AaL and its role in teaching and learning of L2 writing. Lam (2018b) avers that AaL is a fuzzy and slippery construct when it comes to
understanding its definition. AaL could be a subset of AfL; or it is perhaps regarded as a plethora of learner-centered educational principles, which demand a new understanding of self-assessment, self-regulated learning, self-reflection, autonomy, and motivation in L2 writing. While it is not at all easy to agree upon a unitary definition of AaL, scholars, at least, develop a set of down-to-earth and sustainable principles to describe what AaL is, why it is pedagogically significant, how it is best applied in L1 and L2 writing classrooms, and what impact it typically has on teachers and learners. Without having a clear-cut working definition of AaL, researchers may not easily select the most appropriate methods for empirical investigations.

Renewed Interpretation of Assessment Principles

Investigating AaL in large-N studies may be labor-intensive and difficult to manipulate multiple variables in experimental settings. To this end, applying the positivist notions of validity, reliability, and generalisability to AaL studies is perhaps misguided, given that these studies are usually contextually unique, small-scale, exploratory, and longitudinal by nature. Also, they primarily occur in the classroom, implying that the findings are likely to deepen the researcher’s or teachers’ understanding more than making a wider statistical generalization. Because of their small sample size and classroom-based nature, AaL studies may not fulfill the attributes of validity, reliability, and generalizability in a psychometric sense. Instead, AaL researchers could adopt the ideas of ecological validity, dependability, and theoretical generalizability to make sense of the findings of these AaL studies. If not developing a renewed interpretation of these assessment principles under the qualitative research paradigm, researchers may find it problematic to conduct AaL studies rigorously.

Use of Technology in Research Methods

As previously discussed, the scrolling back method on Facebook Timelines is proven to be an option for qualitative researchers who want to find out informants’ lived experiences and learning histories when they grow up. In AaL, the use of electronic narrative frames, weblogs, digital portfolios, social media platforms (e.g., Facebook and Instagram) may facilitate students to monitor their learning trajectories effectively. First, these digital media are accessible as long as there is Wi-Fi connection and availability of smart phones. Second, young adolescents are generally more motivated to use digital tools than paper notebooks when recording their metacognitive composing strategies (Hanna, 2012). Third, the inclusion of Web 2.0 technology in qualitative research methods is time-saving for data input and accordingly facilitative of exporting data to qualitative data analysis software, namely NVivo or MAXQDA 2018. Using digital media to research AaL in L2 writing may be a brand-new endeavor, yet the multimodal nature of these tools could capture the subtlety of metacognitive thinking and reflexive composing processes more accurately.

Provision of Training in Researching AaL

In qualitative research studies, researchers are probably familiar with techniques like interviews, questionnaires, or observations. However, they may not have a full understanding of how narrative frames, verbal reports, draw-and-talk techniques, and social media are used as proper research methods. Thus, teacher trainers, applied linguistics instructors, and professors may consider incorporating alternative research methods (e.g., Skype, WeChat) into their research methodology courses because their postgraduate students may need the training to research AaL or similar complex constructs, which cannot be easily identified by self-reported data. Further, as qualitative researchers, we should be open-minded and receptive to all feasible methods, which are best suited to research AaL, for instance, retrospective verbal reports. This method is valid to explore a range of self-regulated learning and composing processes. Yet researchers need the training to conceptualize the rationale behind this method and additional practices in performing the stimulated recall tasks with their informants.

In conclusion, this article has achieved its aims by illustrating what AaL entails and its role in L2 writing. More importantly, this article discussed how to research AaL from a qualitative research perspective by critiquing five AaL studies and suggesting four alternative research methods to examine AaL, namely narrative frames, verbal reports, draw-and-talk techniques, and social media. The critical review and pilot study have highlighted the potential need to adapt these alternative methods recognized in the literature to suit specific educational contexts, characteristics of learners, and their developmental status. This article is likely to pave the way for further empirical investigations of how the four qualitative methods are validated. Based upon our discussion, we notice that AaL in the L2 writing context is researchable by a range of innovative qualitative methods (e.g., narrative, visual, or e-research methods), and they may also yield trustworthy findings so long as these research methods are theoretically proven (Lam, 2018b). Back to the beginning, we argue that AaL is best researched by qualitative research methods owing to its reflexive, process-oriented, context-sensitive, and longitudinal nature. As Hyland (2015) pointed out that writer-oriented research (i.e., AaL) primarily takes place in naturally occurring contexts over an extended period of time, using the experimental approach to research AaL may not be methodologically justifiable. Despite this, we have no intention to downplay the role of quantitative research methods or mixed methods, which can measure the impact of AaL on pupils’ writing achievements through statistical calculation. No matter which research paradigm scholars adopt, they should formulate a feasible research design, which addresses their research agendas practically and dependably. Lastly, AaL in writing can be investigated by qualitative research methods that are congruent with the features of its working definition.
Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the Language Fund under Research and Development Projects 2018–19 of the Standing Committee on Language Education and Research (SCOLAR), Hong Kong SAR (EDB(LE)/P&R/EL/175/4).

ORCID ID
Ricky Lam https://orcid.org/0000-0003-2903-0425

References
Allal, L. (2019). Assessment and the co-regulation of learning in the classroom. Assessment in Education: Principles, Policy & Practice, https://doi.org/10.1080/0969594X.2019.1609411
Andrade, H., & Brookhart, S. M. (2016). The role of classroom assessment in supporting self-regulated learning. In D. Laveault & L. Allal (Eds.), Assessment for learning: Meeting the challenge of implementation (pp. 293–309). Springer.
Barkhuizen, G. (2011). Narrative knowledge in TESOL. TESOL Quarterly, 45(3), 391–414.
Barkhuizen, G. (2014). Revisiting narrative frames: An instrument for investigating language teaching and learning. System, 47, 12–27.
Barkhuizen, G. (2015). Narrative inquiry. In B. Paltridge & A. Phakiti (Eds.), Research methods in applied linguistics: A practical resource (pp. 169–185). Bloomsbury.
Barkhuizen, G., Benson, P., & Chik, A. (2014). Narrative inquiry in language teaching and learning research. Routledge.
Barkhuizen, G., & Wette, R. (2008). Narrative frames for investigating the experiences of language teachers. System, 36(3), 372–387.
Bennett, R. E. (2011). Formative assessment: A critical review. Assessment in Education: Principles, Policy & Practice, 18(1), 5–25.
Bennett, R. E., & Gitomer, D. H. (2009). Transforming K–12 assessment: Integrating accountability testing, formative assessment and professional support. In C. Wyatt-Smith & J. Cumming (Eds.), Educational assessment in the 21st century (pp. 43–61). Springer.
Benson, P. (2013). Narrative writing as method: Second language identity development in study abroad. In G. Barkhuizen (Ed.), Narrative research in applied linguistics (pp. 244–263). Cambridge University Press.
Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. Assessment in Education: Principles, Policy & Practice, 25(6), 551–575.
Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40.
Bruner, J. S. (1983). Child’s talk. Norton.
Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. Language Testing, 27(1), 5–31.
Carless, D., & Lam, R. (2014). The examined life: Perspectives of lower primary school students in Hong Kong. Education 3–13: International Journal of Primary, Elementary and Early Years Education, 42(3), 313–329.
Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. Educational Psychology Review, 24(2), 205–249.
Cox, S. (2005). Intention and meaning in young children’s drawing. International Journal of Art & Design Education, 24(2), 115–125.
Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). Sage.
Curriculum Development Council. (2017). English language education: Key learning area curriculum guide (Primary 1–Secondary 6). The Education Bureau.
Davidson, J., Paulus, T., & Jackson, K. (2016). Speculating on the future of digital tools for qualitative research. Qualitative Inquiry, 22(7), 606–610.
Davies, A., Pantzopoulos, V., & Gray, K. (2011). Emphasising assessment “as” learning by assessing wiki writing assignments collaboratively and publicly online. Australasian Journal of Educational Technology, 27(5), 798–812.
Duguay, S. (2014). “He has a way gayer Facebook than I do”: Investigating sexual identity disclosure and context collapse on a social networking site. New Media & Society, 16(6), 891–907.
Earl, L. M. (2013). Assessment as learning: Using classroom assessment to maximize student learning (2nd ed.). Corwin.
Earl, L. M., & Katz, S. (2008). Getting to the core of learning: Using assessment for self-monitoring and self-regulation. In S. Swaffield (Ed.), Unlocking assessment: Understanding for reflection and application (pp. 90–104). Routledge.
Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis: Verbal report as data. The MIT Press.
Feltham, M., & Sharen, C. (2015). “What do you mean I wrote a C paper?” Writing, revision, and self-regulation. Collected Essays on Learning and Teaching, 8, 111–138.
Gehl, R. W. (2014). Reverse engineering social media. Temple University Press.
Gottlieb, M. (2016). Assessing English language learners: Bridges to educational equity (2nd ed.). Corwin.
Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. Contemporary Educational Psychology, 34(1), 18–29.
Hanna, P. (2012). Using internet technologies (such as Skype) as a research medium: A research note. Qualitative Research, 12(2), 239–242.
Harris, L. R., Brown, G. T. L., & Harnett, J. A. (2009). “Drawing” out student conceptions: Using pupils’ pictures to examine their conceptions of assessment. In D. M. McNerney, G. T. L. Brown, & G. A. D. Liem (Eds.), Student perspectives on assessment: What students can tell us about assessment for learning (pp. 53–84). Information Age Publishing.
Hawe, E., & Dixon, H. (2014). Building students’ evaluative and productive expertise in the writing classroom. Assessing Writing, 19, 66–79.
Heritage, M. (2010). Formative assessment: Making it happen in the classroom. Corwin.

Holliday, E. L., Harrison, L. J., & McLeod, S. (2009). Listening to children with communication impairment talking through their drawings. Journal of Early Childhood Research, 7(3), 244–263.

Hu, J., & Gao, X. A. (2017). Using think-aloud protocol in self-regulated reading research. Educational Research Review, 22, 181–193.

Huang, S. C. (2015). Setting writing revision goals after assessment for learning. Language Assessment Quarterly, 12(4), 363–385.

Hyland, K. (2015). Researching writing. In B. Paltridge & A. Phakiti (Eds.), Research methods in applied linguistics: A practical resource (pp. 335–348). Bloomsbury.

Hyland, K. (2016). Methods and methodologies in second language writing research. System, 59, 116–125.

Lam, R. (2015). Understanding EFL students’ development of self-regulated learning in a process-oriented writing course. TESOL Journal, 6(3), 527–553.

Lam, R. (2018a). Testing, drilling and learning: What purpose does the grade 3 territory-wide system assessment serve? Asia Pacific Education Review, 19(3), 363–374.

Lam, R. (2018b). Understanding assessment as learning in writing classrooms: The case of portfolio assessment. Iranian Journal of Language Teaching Research, 6(3), 19–36.

Lee, C. (2020). A study of adolescent English learners’ cognitive engagement in writing while using an automated content feedback program. Computer Assisted Language Learning, 33(1–2), 26–57.

Lee, I. (2016). Putting students at the centre of classroom L2 writing assessment. The Canadian Modern Language Review, 72(2), 258–280.

Lee, I. (2017). Classroom writing assessment and feedback in L2 school contexts. Springer.

Mak, P., & Wong, K. (2018). Self-regulation through portfolio assessment in writing classrooms. ELT Journal, 72(1), 49–61.

Nguyen, T. C. L., & Gu, Y. (2013). Strategy-based instruction: A learner-focused approach to developing learner autonomy. Language Teaching Research, 17(1), 9–30.

Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes: Towards a wider conceptual field. Assessment in Education: Principles, Policy & Practice, 5(1), 85–102.

Punch, K. (2014). Introduction to social research: Quantitative and qualitative approaches (3rd ed.). Sage.

Robards, B., & Lincoln, S. (2017). Uncovering longitudinal life narratives: Scrolling back on Facebook. Qualitative Research, 17(6), 715–730.

Sadeghi, K., & Rahmati, T. (2017). Integrating assessment as, for, and of learning in a large-scale exam preparation course. Assessing Writing, 34, 50–61.

Tay-Lim, J., & Lim, S. (2013). Privileging younger children’s voices in research: Use of drawings and a co-construction process. International Journal of Qualitative Methods, 12(1), 65–83.

Wiles, R., Crow, G., & Pain, H. (2011). Innovation in qualitative research methods: A narrative review. Qualitative Research, 11(5), 587–604.

Wiliam, D. (2011). What is assessment for learning? Studies in Educational Evaluation, 37, 3–14.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), The future of assessment: Shaping teaching and learning (pp. 53–82). Lawrence Erlbaum Associates.

Wong, A. T. Y. (2005). Writers’ mental representations of the intended audience and of the rhetorical purpose for writing and the strategies that they employed when they composed. System, 33, 29–47.

Xiang, W. (2004). Encouraging self-monitoring in writing by Chinese students. ELT Journal, 58(3), 238–246.

Zimmer, M. (2010). “But the data is already public”: On the ethics of research in Facebook. Ethics and Information Technology, 12(4), 313–325.