An Overview of Writing Process Research: Towards a Better Understanding of L2 Writing Process

Mahshad Davoodifard[1]
Teachers College, Columbia University

ABSTRACT

Over the past 40 years, second language educators and assessors have come to the realization that investigating the process of writing can shed light on language teaching, learning and assessment practices (Odendahl & Deane, 2018). What L2 writers do and think while writing can provide links between the task, the related construct and the cognitive mechanisms the learners engage in. While from a pedagogical perspective, such information can help language teachers provide learners with more accurate feedback, in assessment contexts, writing process information can help language testers to ensure that performance-based score inferences are truly linked to not only the linguistic knowledge of examinees, but also to the functions of mental architecture intended in a specific task (Purpura, 2013). Despite the abundance of writing process research in L1 and L2, most studies suffice to a mere list of strategies and the observed phenomena (Manchón et al., 2007). In light of the demand for a theoretically-based approach to L2 writing process, this paper first provides a review of L2 writing process research by presenting the most influential theoretical approaches to the conceptualization and investigation of writing process in L1 and L2. Then it evaluates the contributions made to the concept of writing process by reviewing some representative empirical research in the field and delves more deeply into the investigation of writing process in assessment contexts. Finally, it discusses how a stronger, theory-based understanding of L2 writing process can be achieved by means of using more innovative assessment tools such as scenario-based assessment (SBA) in conjunction with computer technology.

Keywords: cognition, metacognition, scenario-based assessment, technology, writing strategies, writing process

INTRODUCTION

---

[1] Mahshad Davoodifard is a doctoral student in the applied linguistics program at Teachers College, Columbia University. She is currently the assessment coordinator and a doctoral fellow at the Community Language Program. After completing her postgraduate research and studies at Monash University, Australia, Mahshad Joined Educational Testing Service in 2012, and worked as a test developer and professional rater of the spoken Test of English for International Communication. Mahshad started her doctoral work at TC in 2017 under the supervision of Professor James Purpura. Her research focuses on the use of technology for better understanding and enhancement of second language assessment. Correspondence should be sent to E-mail: md3573@tc.columbia.edu

Studies of second language (L2) writing performance and development are taking new directions to respond to findings from research in writing ability and cognitive theories (Deane, 2011; Michel et al., 2020; Roca de Larios et al., 2008). Among the existing approaches to L2 writing research, the process approach started to develop in the early 1980's as a result of a shift of attention from the writing product to the writer and the writing process (Krapels, 1990a). Unlike the product-oriented approach to writing, which focuses on the final product of the writing task, the process-oriented approach considers not only the linguistic aspects of writing (e.g., knowledge of grammar), but also the cognitive and metacognitive aspects that involve the necessary phases a student goes through to produce a text. These phases can include brainstorming, planning, generating ideas, and revising and reflecting on one's own work (Odendahl & Deane, 2018). In the literature, different terms have been employed to explain the process of writing (e.g., strategies, aspects, processes, behaviors). For the purpose of this paper, writing process refers to the combination of strategies used by language learners to produce text. This includes both the strategies that have a clear physical representation (e.g., writing words on a paper), as well as the underlying cognitive and metacognitive processes that guide and form the actual witting process.

In L2 assessment contexts, as Purpura (2013) points out, language testers have been interested in researching the cognitive mechanisms underlying L2 performance to establish a link between the assessment tasks, the related construct and the mental processes that examinees engage in during assessment. It is also critical for language testers to assure the score-based inferences depend on both the linguistic knowledge of the examinees and the functions of their mental cognitive architecture (e.g., working and long-term memory) in their performance (Purpura, 2013, p. 1453). Being a complex cognitive ability, writing can pose additional affective and cognitive demands on language learners when performing under assessment conditions (Deane et al., 2008). Investigating writing process in assessment contexts, as such, can provide valuable information about the underlying mental processes leading to the final product (Choi & Deane, 2021).

Despite the importance and value in discovering aspects of L2 writing process, research into L2 writing process in assessment contexts are rather scarce. Recognizing the challenges involved in the investigation of L2 writing process in assessment contexts, the purpose of this paper is to review the literature on L2 writing process by first presenting the theoretical models of writing process, then reviewing the methodology and major findings of representative L2 empirical research based on early theoretical frameworks, and finally examining the more recent L2 writing process research in assessment contexts.

THE THEORETICAL MODELS OF WRITING PROCESS

Second language writing research has made extensive use of L1 process-oriented models of writing (e.g., Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; Hayes, 1996; Hayes, 2012; Kellogg, 2001), which describe the mental processes through which native language writers compose texts (Cumming, 2016; Roca de Larios et al., 2016; Silva & Leki, 2004). One of the seminal L1 process-oriented models of writing was proposed by Flower and Hayes (1980). In this framework, the writing process is divided into three subprocesses of planning, translating (of ideas into words), and reviewing, all controlled by the process of monitoring. When engaged in

writing, these subprocesses interact in a recursive manner and work hand in hand with task environment, control level (including the task goals) and resource level (e.g., long-term memory). The Hayes-Flower framework has been criticized, however, for not accounting for the language learners' linguistic knowledge as well as the external, social factors affecting the writing process (Deane et al., 2008; Weigle, 2002). Nevertheless, the Hayes-Flower model has been used in a large number of L2 process studies, and as Hayes (2012) pointed out, several main features of the framework are still valid in the modern representations of writing.

Another influential model of writing process was proposed by Bereiter and Scardamalia (1987), who suggested that people may follow two models of composing: knowledge telling and knowledge transforming. According to Bereiter and Scardamalia, knowledge telling lets people use their maximum language competence and skills through the natural social experience. Knowledge transforming, on the other hand, requires the individual to go beyond their normal linguistic competence and reprocess knowledge independent from ordinary social interaction. Bereiter and Scardamalia (1987) argued that beginner writers use the knowledge telling model to retrieve appropriate content from memory and speedily produce discourse. The advanced writers, on the other hand, use knowledge transforming strategies to not only express their thoughts, but also to develop their knowledge by constantly reviewing what they wrote and what they want to say. Bereiter and Scardamalia's (1987) model was not used as widely as the Hayes-Flower model in L2 studies, since L2 writers' strategies were found to be different from and not as developed as L1 writers described in this framework (di Gennaro, 2006). For example, compared to less-skilled native speakers, novice L2 writers may find it less natural and more challenging to use knowledge-telling as a strategy to compose a text. The model was also criticized for not accounting for the underlying process of transformation from knowledge telling to knowledge transforming, making it difficult to determine if a writer's ability was in the middle of the telling vs. transforming continuum (Grabe & Kaplan, 1996).

Adapting a communicative language performance model originally proposed by Chapelle et al. (1993), Grabe and Kaplan (1996), proposed a process-oriented model of writing, which specified the process as comprising both the writer internal processing and external factors. The external social context includes the variables of situation (participants, setting, task, text, and topic), and performance (textual output). The internal processes refer to the processing activities of the writer in verbal working memory. According to Grabe and Kaplan's (1996) model, when given a writing task, the writer sets up goals and purposes based on either the "context" or the "internal goal setting", which in turn activates a cycle of operations in "verbal processing", mediated by metacognitive factors. The writer uses language competence along with their knowledge of the world and produces the internal output, which is compared to the goals of the task and is revised if necessary. As Grabe and Kaplan (1996) point out, the processing sequence in this model is not very different from either the Hayes-Flower (1980) model or Bereiter and Scardamalia's (1987) framework. Grabe and Kaplan's (1996) model is distinguished in terms of how contextual influences, task specifications and different sources of processing are incorporated.

Yet another model of writing, which is frequently cited in L2 writing process research, was proposed by Kellogg (1996). In this model, writing consists of six basic processes: 1) planning, 2) translating, 3) programming, 4) executing, 5) reading, and 6) editing. Kellogg's (1996) model differs from Hayes-Flower model in the working memory demands specified for each subprocess. For example, in Kellogg's model, it is argued that generating and organizing ideas pose the same demands on the working memory. Translating in this model includes the

subprocesses of linguistic encoding (e.g., grammatical, phonological, and orthographic encoding), responsible for translating the activated concepts into words and sentences. In a review of this model, Kellogg et al. (2013) argue that theoretical and empirical research have been supportive of the assumptions made in Kellogg's (1996) model, even though working memory resources called for in the model may differ depending on tasks or type of writing. For instance, abstract concepts do not activate visualization during planning.

Recognizing the need to include contextual and social factors as well as critical thinking abilities as parts of the writing process, Deane et al. (2008) proposed a writing competency model with three basic strands: I) language and literacy skills for writing, II) writing process management skills, and III) critical thinking for writing. This model was designed to fit in an assessment context, while accounting for teaching implications. In Deane et al.'s (2008) model, Strand I comprises of a set of skills such as drafting and editing a text. They are considered automatic since they are related to the individual's ability to use Standard English and basic literacy skills. Strand II includes writing-process management skills such as planning and evaluating a document. It also includes the ability to manage the drafting process to write a well-developed and well-organized text. Strand III refers to critical thinking abilities enabling the individual to reason about the content as well as the social context when writing. In terms of content, these abilities let the writer explain, hypothesize, support, refute, synthesize, report, narrate or describe. In terms of the social context, the individual makes decisions regarding collaboration and review, hence fulfilling a social role, or thinks about the audience and how to engage them. Furthermore, as Deane et al. (2008) explain, the elements of this model can vary based on the requirements of the specific genre and task and allow for different combination of skills and strategies. This comprehensive model is therefore capable of being applied to investigation of L2 writing process in innovative tests such as scenario-based assessments.

Table 1 summarizes the major components of these influential theoretical models of writing process. The major theoretical frameworks presented in this section outline the cognitive, metacognitive, and physical processes that writers go through to produce text. Although the major components of the writing process (planning, generating ideas, writing, organizing, and editing) are shared among different models, they differ in their emphasis on the linguistic, social, and contextual features. What is evident from all approaches is that writing involves complex cognitive mechanisms and problem-solving abilities. In the following section, a review of the general L2 writing process research influenced by the major theoretical writing process frameworks are presented and reviewed.

TABLE 1
Summary of the Early Theoretical Frameworks of Writing Process

Author(s)	Year	Components of the Writing Process

Flower & Hayes	1980	• Planning
		    o Organizing
		    o Goal setting
		• Translating
		• Reviewing
		    o Reading
		    o Editing

Bereiter & Scardamalia	1987	• Knowledge telling
		    o Content knowledge
		    o Discourse knowledge

		- o Topic and genre identifiers - o Construct probes and retrieve memory - o Test for appropriateness - o Write - Knowledge transforming - o Problem analysis and goal setting - o Content knowledge - o Discourse knowledge - o Content and rhetorical problem space - o Write and reflect
Grabe & Kaplan	1996	- External social context - o Situation (participants, setting, task, topic, register and genre constraints, communication purpose, language use norms) - o Performance (textual output) - Internal processing (Verbal processing memory – Metacognitive processing) - o Internal goal setting - o Language competence (linguistics, sociolinguistic, discourse) - o Knowledge of the world - o On-line processing assembly - o Internal processing output
Kellogg	1996	- Planning - Translating (Linguistic encoding) - Programming - Executing - Reading - Editing
Deane, Odendahl, Quinlan, Fowles, Welsh, Bivens-Tatum	2008	- Strand 1 - o Automatic processes (language and literacy skills for writing) - Strand 2 - o Text generation - o Organization - o Evaluation - Strand 3 - o Critical thinking abilities ▪ Content (rhetorical goals) ▪ Social context (social roles - audience)

INVESTIGATING L2 WRITING PROCESS: L2 EMPIRICAL STUDIES

The investigation of L2 writing process is believed to be motivated by pedagogical purposes (Roca de Larios et al., 2002), assuming that it could provide teaching practices with a sound theory of writing. As a result, many L2 writing process studies investigated the effectiveness of L1 process theories for L2 writing instruction (Cumming, 1989; Silva, 1993; Zamel, 1982). Influenced by major models of writing described by Flower and Hayes (1980) and Bereiter and Scardamalia's (1987), many of these studies based their arguments on the differences between L1 and L2, paying attention to skilled vs. unskilled writers, as well as effects of fluency (e.g., Chenoweth & Hayes, 2001). Some studies investigated writing process at a global level (e.g., Raimes, 1987) or explored specific sub-processes such as planning or revision (e.g., Cumming, 1989; Krapels, 1990b; Roca de Larios et al., 2001; Roca de Larios et al., 1999), or examined differences in time allocation (Roca de Larios et al., 2008). Some other studies took a developmental approach and examined L2 writers' development of writing strategies over time (e.g., Sasaki, 2002).

The studies conducted by Zamel (1982, 1983) and Raimes (1985, 1987) are among the first ones investigating L2 writing process. Primarily concerned with improving writing instruction, Zamel (1983) examined the composing process of skilled and unskilled ESL students by observing their writing behavior as they wrote several drafts of a formal university-level expository essay and interviewing them at the end of the composition. Zamel argued that thinking, writing, and rewriting were the strategies used in a recursive manner by ESL students. Unskilled writers in Zamel's study were caught up in correcting surface grammatical and lexical errors compared to skilled writers, who were mostly concerned with developing their ideas more sufficiently, similar to native speakers. Zamel (1983) concluded that writing is a process of discovering ideas and finding the best framework to present and develop these ideas. Raimes (1987) found the skilled-unskilled distinction in Zamel's study blurry and argued that such a definition can be applied to ESL learners only after considering their level of language as well as their L1 writing ability. Using two different tasks and think-aloud protocols and retrospective interviews, Raimes (1987) examined writing process of ESL students at different levels of instruction (remedial vs. college-level ESL writing courses). Finding no correlation between language proficiency and writing strategies, Raime reported that lower L2 proficiency level students acted similarly to less skilled L1 writers in terms of the short time spent on prewriting and planning. The more advanced L2 students, on the other hand, spent more time reading and revising. Raimes's (1987) replicable research procedures and the use of different tasks and proficiency levels turned the study into a valuable contribution to the field of L2 writing research, despite the limitations due to a relatively small sample size and data collection method (di Gennaro, 2006).

Cumming (1989) agreed with Zamel (1983) and Raimes's (1987) findings indicating that some L2 writing processes are comparable to L1, but questioned the claims made by those studies regarding the lack of effect of L2 proficiency on writing process in L2. Cumming (1989) argued that L2 proficiency and writing expertise must be treated as separate variables, and L2 process studies should also examine the effect of tasks more systematically. To address these concerns, Cumming (1989) conducted a larger scale study on L2 writing process by investigating the writings and think-aloud data from college students responding to three different tasks (letter writing, summary, argumentation). Cumming found that students from higher L2 proficiency and writing expertise scored higher than lower-level students, with the difference being more

apparent for more cognitively demanding tasks. Moreover, while almost all students attended to the gist, or substantive content of their writing across all tasks when thinking, the experts attended to word-level evaluations after making decision about the gist and organization of their texts. In contrast, the nonexpert writers were constrained in formulating their gist and focused frequently on immediate planning and correction. Cumming (1989) also reported that the students applied their L1 writing expertise in their L2 writing, and L2 proficiency level did not necessarily affect the quality of writing process and problem-solving behaviors when writing, a finding that lent support to previous research in the field.

In terms of methodology, while verbal protocols used in these studies generated valuable information, their validity was questioned for causing reactivity and veridicality (Janssen et al., 1996; Stratman & Hamp-Lyons, 1994). This led researchers to use less obtrusive methods such as computer programs and video taping writing sessions. Sasaki (2002), for example, examined novice and expert Japanese L2 learners of English expository writing process and asked them to explain their writing behavior while watching their tape-recorded composition sessions. Sasaki's (2002) study showed that expert writers spent more time planning and produced lengthier texts in a shorter time compared to their novice counterparts. Sasaki (2002) also found that novice writers spent more time on local planning as well as translating from their L1 into L2 due to low L2 proficiency. The experts on the other hand, spent more time planning the content as well as the best way to present their ideas, a finding that again mirrored the trends in previous research.

A concurrent line of research started to examine the context of writing and the effect of a set time limit on L2 writing process. Research findings suggested that L2 learners' process and subprocesses differ when performing under assessment conditions compared to non-assessment settings (e.g., Ellis & Yuan 2004; Hall, 1991). Khuder and Harwood (2015), for example, examined the writing process of ESL students writing an argumentative essay under both a test and a non test situation. Khuder and Harwood found that while translation and surface revision were used frequently in test situation, the students did more planning and evaluation in the non-test condition. In terms of time allocation, overall, the participants paused longer to make surface evaluations in both situations but spent more time for distant meaningful revisions in non-test situation, and immediate sentence revisions in test situation. The role of task type on L2 writing process, particularly in assessment contexts, also started to receive attention among researchers (Cumming et al., 2005; Plakans, 2008). The combined effects of writing context, L2 proficiency, task types and data collection methods on language learners' writing process motivated the study of L2 writing process under assessment conditions. Some of these studies and their major findings are reviewed in the following section. Table 2 summarizes the details of the representative L2 writing process research studies reviewed in this section.

TABLE 2
Summary of Studies on L2 Writing Process

Author(s)/ year	Language and process focus/ Task/ Task conditions	Participants	Technique	Results
Zamel (1983)	• ESL • Global writing process • Task: Formal expository academic essay • Task Conditions: several drafts of the same essay produced under no time limit	University students, skilled and unskilled L2 writers, enrolled in an elective intermediate composition class, different L1 backgrounds (N=6)	• Observation of the writing behavior (observing students as they write) • Recording students' comments while writing • Retrospective interviews	Writing is a non-linear recursive process of discovering meaning and formulating ideas. Skilled writers attend to language-related concerns only after ideas are delineated. Unskilled writers spend more time fixing surface syntactic and lexical errors.
Raimes (1987)	• ESL • Global writing process • 2 tasks: (1) explicit purpose and audience, (2) a standard academic "agree/disagree" topic • Task Conditions: one draft of each essay written with no time limit	University students enrolled in advanced, remedial and college-level ESL writing courses (N=8)	• Think-aloud protocols • Retrospective interviews	No correlation was found between the students' L2 proficiency level and their writing ability and composition strategies. More skilled writers were more engaged with the text by reading and revising. Less skilled writers spent less time prewriting and planning.
Cumming (1989)	• ESL • Writing and problem-solving behavior • 3 tasks (letter, argument, summary) • Task condition: one draft within 3 hours per task	Young adult ESL learners from 3 L1 writing expertise and 2 L2 proficiency levels (N=23)	• Think-aloud protocols	L2 proficiency and L1 writing expertise significantly affected writing and problem-solving behaviors of the students, with expertise accounting for the major discourse and organization affects and L2 proficiency being an additive factor.
Sasaki (2002)	• EFL • Global writing process, examined cross-sectionally and longitudinally • Task: argumentative	Expert (N=12), and novice (N=22) Japanese EFL learners	• Recording of written product • Video-taped writing behavior • Simulated recall protocols	Experts wrote longer texts in shorter time and spent more time planning in detail. Instruction over two weeks did not improve novice's speed or the quantity of writing but improved their planning skills. Experts refined ideas during writing, while novices translated their ideas into L2 or engage in local planning only.

	• Task condition: one prompt given to two groups; another prompt given to novices after 2 weeks			
Khuder & Harwood (2015)	• ESL • Global writing process and time allocation • Task: Argumentative essay from IELTS • Task condition: test (40 mins) vs. non-test (unlimited time, access to internet)	L2 postgraduates (varied L1 backgrounds) (N=10)	• Keystroke logging • Screen recording • Simulated recall protocols	Test situation processes: translation from L1 and surface revisions Non-test situation processes: evaluation and meaningful revisions; essays were scored higher Statistically significant time allocation difference between processes were found.

INVESTIGATING L2 WRITING PROCESS IN ASSESSMENT CONTEXTS

Over the last decade, an increasing number of studies have provided evidence that discovering and understanding writing process in assessment contexts can give us useful information about the cognitive mechanisms underlying performance in writing tasks. Such information is useful in not only assuring that these mechanisms are congruent with the intended construct but that they are also relevant for successful performance (Purpura, 2013). The limiting effects of testing situation on writing process were documented as early as 1980's (Wolcott, 1987), arguing that assessment context can affect the writer's engagement with text. Global and local writing processes have been investigated in assessment contexts in L1 and L2 (Miller, 2000; Worden, 2009) as well as across different task types, including the use of source materials (Barkaoui, 2016; Cumming et al., 2005; Gebril & Plakans, 2013; Révész et al., 2017).

In a study of writing strategy use by ESL students, Yang and Plakans (2012) used a questionnaire to investigate how self-regulatory, discourse synthesis, and test-wiseness strategies were coordinated during a TOEFL iBT integrated writing task. Yang and Plakans found that L2 proficiency did not necessarily lead to a better source use ability and a higher score. However, the writers had to activate a combination of self-regulatory strategies to perform successfully in the task. While test wiseness strategies had a negative effect, discourse synthesis strategies had a significantly positive effect on writing successfully. Yang and Plakans's (2012) study supported the notion of integrated task writing as being a multifaceted construct. Their results, however, were limited by context and method used. Although the task was timed and selected from a high-stakes standardized test, the students responded under research setting and not a real testing situation, which may have influenced the type of strategy use reported. The use of questionnaire could have also limited the students' responses and excluded potential strategies used in specific situations by test takers.

Taking a similar approach, Barkaoui (Barkaoui, 2015), examined the writing activities of test takers responding to TOEFL iBT writing tasks, but unlike Yang and Plakan's (2012) study, Barkaoui (2015) accounted for L2 proficiency (high and low) and task type differences (integrated vs. independent) and used simulated recalls. Barkaoui found that the evaluation of language and local text was the most frequently observed activity in all tasks. Writers interacted with the integrated writing task more often and generated source-based content, but they revised language more frequently in the independent task. This result is expected as participants had no access to extra materials and had to self-generate more content during the independent task. Barkaoui (2015) also found that in the independent task, lower L2 proficiency level students did more planning and organization activities than high L2 participants. They also reported the independent task more difficult than the integrated task. High L2 level participants, on the other hand, reported less interaction but more evaluation and revising activities with the independent task. As Barkaoui (2015, p. 21) points out, these writing activities can be considered part of the writing construct as they are in line with strategies expected when responding to different task types. However, simulated recalls may not fully capture the broad range of writing activities that writers engage in during the test, as they may not precisely recall the process they went through after the task is complete.

With the majority of high-stakes language tests becoming computerized, and given the limitations imposed by verbal protocols and questionnaires, an increasing number of L2 process studies started to rely on new techniques such as keystroke logging programs to examine writing process (Almond et al., 2012; Barkaoui, 2016, 2019; Chukharev-Hudilainen, 2019; Conijn et al.,

2019; Leijten & Van Waes, 2013; Miller et al., 2008; Révész et al., 2017; Zhang et al., 2017). Unlike verbal protocols, keystroke logging enabled researchers to collect writing process data in a non-invasive and natural context. For instance, taking into account some factors that can contribute to difficulty of L2 writing assessment, Révész et al. (2017) used keystroke logs in combination with simulated recalls and a perception questionnaire to examine the effect of task complexity (i.e., presence or absence of content support) on L2 writers' fluency, pausing and revision behavior and the underpinning cognitive processes.

      Révész et al. (2017) found that test takers reported the simple task as requiring less mental effort and being less difficult in terms of content generation. During the simple task, test takers paused less and translated more than the complex task, during which they used more planning strategies. Révész et al. (2017) concluded that simple writing tasks, which provide content support to test takers, would decrease pressure on planning processes, which can lead to more extensive higher order composing processes. The results of Révész et al.'s (2017) study, however, are limited by the single task used and L2 proficiency level of the participants, who were all at advanced level. It would be informative to examine how low L2 level participants approach different writing tasks from a cognitive and process-oriented point of view.

      To account for the effects of different task types and proficiency levels, Barkaoui (Barkaoui, 2019) investigated the pausing patterns of ESL students responding to two TOEFL iBT integrated and independent tasks. By analyzing students' keystroke logs, Barkaoui (2019) found that the participants completed the integrated task faster than the independent task, but L2 proficiency had no effect on the overall pause duration. The frequency of pausing, however, was higher for low L2 participants. Higher proficiency L2 students paused more at the beginning to read the task and plan. Barkaoui also found that students paused more between paragraphs in the independent task, probably to plan more. In terms of location, there were more pauses between and within words in both tasks. While Barkaoui's (2019) findings were in line with previous research, the participants were not asked to comment on their pausing behavior, which may have influenced the interpretation of the pause patterns. In other words, keystroke patterns were recorded and associated with what was assumed to be the true reasons accounting for those patterns. As Galbraith and Baaijen (2019) point out, keystrokes need to be aligned with cognitive processes, as different patterns of keystrokes (including pauses and bursts) may reflect various levels of writing process underlined by certain cognitive mechanisms. Therefore, keystroke logging needs to be triangulated with other methods, such as retrospective interviews, to provide a clearer idea of the nature of process features and their underlying cognitive mechanisms.

      In a mixed-method study, Michel et al. (2020) used keystroke logging along with eye gaze tracking and simulated recall interviews to investigate the writing behaviors and the associated cognitive processes of Chinese intermediate and advanced level ESL students. The participants completed two research versions of TOEFL iBT independent and integrated writing tasks. Pause patterns and speed fluency were recorded by keystroke logs, and eye-gaze tracking provided the researchers with the students' possible source use during the assessment. The participants' cognitive processes were obtained from simulated recall interviews where the participants were presented with episodes of their composition processes (recordings of screen while writing) and asked to comment on their pausing patterns. Michel et al. (2020) found that task type had a significant effect on the writers' speed and fluency. They had less frequent and shorter pauses in the independent task since they did not have to refer to notes from the listening and reading sections, like they did in the integrated task. Michel et al. also found that the pause patterns changed based on the stage of writing in the integrated task. For example, the students

paused more frequently during the first and the last stages of writing where they planned and monitored their essays. For the independent task, on the other hand, these differences were observed between the first stage of writing and the middle sections, where most students revised their essay at word-level. With respect to cognitive mechanisms, Michel et al. (2020) found that in both task types students relied heavily on translation, but they spent more time planning during the independent task, which was due to lack of access to source materials. The results from Michel et al.'s (2020) study confirms the findings from previous research on integrated tasks (e.g., Barkaoui, 2016), reporting source use at initial stages, attention to and revisions of content in the middle stages, and monitoring and revisions in the final stages. Their results are also in line with Révész et al.'s (2017) study, indicating that the content support provided in integrated tasks can decrease planning processes and allow the writers to concentrate on other processes such as overall revision and monitoring. However, similar to many other process studies, their findings are limited to the specific context of the study, which is mostly due to the type of tasks used. The majority of L2 writing process studies used TOEFL iBT integrated and independent writing tasks and ignored the social variables involved in real-life writing situations, such as the purpose of the task, the role of the writer and the audience. They also did not replicate a true assessment condition. According to Roca de Larios et al. (2002, p. 45), Michel et al.'s study of L2 writing process may be limited based on the contextual and social factors as well as factors affecting students' performance under assessment conditions. Furthermore, Michel et al. (2020) based their argument on an L1 model of writing (Kellogg's 1996 model), disregarding necessary second language linguistic knowledge and skills, and the cognitive demands L2 can impose on writers, particularly in assessment contexts.

Taking a different approach to L2 writing process in assessment contexts, Choi and Deane (2021) explored the stability and predictive potential of writing process features of adult EFL learners. The participants completed two tasks, an integrated source-based writing, and an online discussion situated in an online forum format eliciting a written argument. The goal of Choi and Deane's study was to discover the most likely L2 writing processes associated with participants' pause behaviors and examine how stable those features were as the writers neared the completion of the task. Choi and Deane (2021) also built a model for predicting human rater scores based on the most stable writing process features. They found that only stable features (such as total keystrokes and within-word pauses) correlated with the quality of the response and overall English proficiency. In other words, test takers with a higher number of logged keystrokes, consistent within-burst pauses, and shorter within-word pauses scored higher by human raters. Moreover, the amount of writing, writing fluency and variance in pause duration correlated with one another, supporting previous research findings. Choi and Deane (2021) also found a consistent relationship between potential indicators of students' keyboarding skills and human scores. Although the findings of this study are limited by its exploratory nature and specific context, they are informative to language test researchers due to the kinds of tasks used, and their approach to evaluating writing process features. Unlike past studies, Choi and Deane's (2021) study accounted for the specific purposes and audiences for each writing task (social variables) and treated pauses in writing within their context. However, similar to previous studies (e.g., Barkaoui, 2019), they relied on keystroke logs and the strongest association was assumed between the writing patterns and cognitive mechanisms.

The studies conducted on L2 writing process in assessment contexts (see Table 3) suggest that assessment conditions affect a variety of writing behaviors and cognitive mechanisms such as time allocation, pausing, planning, monitoring, fluency, revision patterns and perception of

difficulty level. As the second column in Table 3 shows, most writing process studies in assessment contexts used samples of high-stake tests such as TOEFL iBT and showed that writing processes are also affected by the type of task and having access to content support or source materials. For example, integrated task types can be easier for higher L2 proficiency level writers by reducing planning processes and allowing them to focus on higher order processes such as global monitoring and revisions. Lower L2 proficiency level writers, however, may find integrated tasks harder to process and use. The effects of L2 proficiency level on writing process were further evidenced in studies showing that high proficiency students normally engaged in more planning, rescanning, revising, and editing. Less proficient students, on the other hand, had less diversified allocation of time, and did more local planning and corrections (Hall, 1991; Roca de Larios et al., 2008; Sasaki, 2002).

TABLE 3
Summary of Studies on L2 Writing Process in Assessment Contexts

Author(s)/ year	Language focus/ Task/ Task conditions	Participants	Technique	Study focus	Results
Yang and Plakans (2012)	• ESL • Writing strategy use • Task: an integrated listening-reading-writing task from TOEFL iBT • Task condition: 20 minutes to plan and write 150 to 225 words, complete questionnaire immediately after	University students, varied L1 backgrounds, (N=161)	Integrated writing task, Strategy use questionnaire	• Writers' coordination of self regulatory, discourse synthesis and test wiseness strategies	A combination of self-regulatory strategies is needed for successful performance on integrated tasks. L2 proficiency was not related to source use ability. There was a positive effect of discourse synthesis strategies and a negative effect of test wiseness on performance.
Barkaoui (2015)	• ESL • Global writing activities • Task: 1 integrated and 2 independent tasks from TOEFL iBT • Task condition: timed, one independent task paper based, the other tasks computer based	High and low proficiency level and keyboarding skill, university and ESL class students (N=97)	Video-taped writing sessions, Simulated recalls	• Effects of L2 proficiency level, keyboarding skills, task type on writing activities	All participants reported interacting with source materials, followed by planning and organizing in integrated task. They revised more in independent task. Low L2 students found integrated task more difficult and did more planning on independent task. Advanced L2 students engaged with independent task less and did more evaluation and revising in it.
Révész, Kourtali, Mazgutova (2017)	• ESL • Fluency, pausing and revision behavior • Task: argumentative essay • Task condition: Students randomly assigned to simple and complex versions of the same task. 8 students selected for simulated recalls. Writings	Advanced ESL International university students (N=73)	Keystroke logs, Perception questionnaire, Simulated recalls	• Effect of task complexity (presence vs. absence of content support) on writing processes	Students reported the simple task to be easier, requiring less mental effort. They paused less and translated more in simple task but used more planning strategies when no content support was given.

	done in a quiet computer room.				
Barkaoui (2019)	• ESL • Pausing behavior • Task: TOEFL iBT integrated and independent tasks (1 each) • Task condition: timed, one draft of each task	High and low proficiency level and keyboarding skill, university and ESL class students (N=68)	Keystroke logs	• Effects of L2 proficiency level, keyboarding skills, task type on pausing behavior	It took longer for the students to complete the independent task due to lack of access to source materials and the need to plan more. Low L2 proficiency level students paused more frequently to attend to immediate corrections. High L2 proficiency students paused longer at the beginning to read tasks and plan. Task type had no significant effect on overall pause duration.
Michel, Révész, Lu, Kourtali, Lee & Borges (2020)	• ESL • Global writing and cognitive processes • Task: TOEFL iBT integrated and independent tasks (2 each) • Task condition: one draft of each task (60-70 mins) followed by interview	Chinese ESL university students (N=60)	Keystroke logs, eye-gaze tracking, simulated recall protocols	• Effects of task type on writing processes as a whole and at different writing stages • Writing speed fluency, pausing, eye-gaze behavior, cognitive processes	More varied writing behaviors were found during the integrated task. Task type had a significant effect on speed fluency. Students paused less and for shorter times during independent task since they did not have to look at source materials. They paused for planning and monitoring during the integrated task but paused during planning and middle sections of the independent task for word-level revisions. Students used translation during both task types.
Choi & Deane (2021)	• EFL • Stability and predictive potential of process features • Task: Source-based writing and online discussion • Task condition: students wrote 1 draft of each task under a simulated assessment condition	Adult EFL learners, varied L1, same L2 proficiency level (N=380 completing both tasks)	Keystroke logs	• Stability of writing processes towards the end of composition • Creation of an AI model for predicting human rater scores	Only stable features (total keystrokes, shorter within-burst pause, shorter within-word pause) correlated with overall writing quality and proficiency. Writing length, fluency and variance in pause duration were correlated, indicating that high performers write more in shorter time, pause less and achieve higher scores.

TOWARDS A BETTER UNDERSTANDING OF L2 WRITING PROCESS: INNOVATIVE ASSESSMENT TOOLS AND TECHNOLOGIES

As discussed earlier, an investigation of L2 writing process in assessment contexts has a number of benefits for not only language learners and educators but also language assessors. For example, what L2 writing process in assessment contexts has revealed so far can be used to provide feedback to language learners as well as teachers as to what areas to work on to improve performance in a test. The information can also be used to design writing tests that reflect the use of expected competencies during a task.

One way to accomplish this goal is to use a scenario-based assessment (SBA) of writing. SBA can provide the opportunity to investigate L2 writing process and strategies in an integrated, more user-friendly assessment design (Deane et al., 2015). With its use of scaffolding and thematically related scenarios, SBA is capable of simulating an authentic, purposeful situation where students write to achieve a meaningful objective which ultimately leads to learning in addition to assessment. Furthermore, SBA takes into account indicators of other skills necessary to accomplish a writing task successfully (e.g., reading comprehension).

What language learners think about and actually do during an SBA can further inform test validation by providing evidence about the process of responding and the test design (Messick, 1989; Mislevy, 2007). Since SBAs are fully computerized, keystroke logging can be used in combination with other tools such as simulated recall interviews and natural language processing techniques to identify not only the writing process features relevant to SBA (e.g., source use, planning, revision, text production), but also identify the errors and linguistic complexity of the final text produced by language learners and how they correspond to the overall writing process and the mental processes the leaners engaged in during the assessment. This line of research can offer promising perspectives for the conceptualization of L2 writing process across different proficiency levels as well as validity evidence for new test designs.

Although SBA has been used in a number of writing process research studies (e.g., Guo et al., 2020; Zhang et al., 2017), as discussed in the previous section, most of these studies are limited to L1 writing or to a school-aged young adult population. Moreover, similar to other L2 process research studies, the reported lists of observed strategies linked to writing process phenomena have not been clearly linked to aspects of cognition and metacognition. In other words, as Manchón et al. (2007) argue, in order to have multiple levels of analysis and generalization – and a more valid and complete account of the data– the analysis of the writing process needs to be ideally rooted in a theoretical model of cognition and different types of cognitive and metacognitive strategies involved in and associated with the observed categories of what is generally referred to as writing strategies. This can be accomplished by using SBAs as more authentic assessments of writing ability, where L2 writing process can be better captured and analyzed in reference to particular cognitive and metacognitive processes underlying the writing activities. Moreover, including a wider population of adult ESL learners from different backgrounds and L2 proficiency levels can potentially help us achieve an even better understanding of L2 process in assessment contexts.

CONCLUSION

In order to review the literature on L2 writing process, this paper presented an overview of the major theoretical frameworks and approaches to the conceptualization and investigation of writing process. It then presented and discussed the contributions made by some representative empirical research in general writing and in assessment contexts. Finally, it examined how SBA can be used in L2 assessment contexts to link L2 writing process of more varied population of language users to a theoretically grounded model of cognition and help us achieve a more valid analysis of writing process data and a better understanding of the concept of L2 writing process.

Developments in theoretical models and research methodologies have led the researchers to consider the key role of individual and social factors as well as the requirements of specific writing tasks and genres in writing process research (Deane et al., 2008). More authentic task types and simulated assessment conditions are being used along with less intrusive tools such as keystroke logs and other computer software. One shortcoming with previous research is that the majority of L2 writing process studies relied on high-stakes standardized tests (most of which do not represent a real-life writing task) and *assumed* a relationship between the observable writing strategies and cognitive and metacognitive processes underlying these strategies. It is, therefore, essential to address the issues and implications emerging from the past and current L2 writing process research by investigating the writing behaviors and underlying cognitive mechanisms using innovative types of assessment that consider the wide range of contextual factors, and personal and social knowledge, skills and abilities needed to successfully fulfill a writing task.

REFERENCES

Almond, R., Deane, P., Quinlan, T., Wagner, M., & Sydorenko, T. (2012). A preliminary analysis of keystroke log data from a timed writing task. *ETS Research Report Series*, *2012*(2), i–61.

Barkaoui, K. (2015). Test takers' writing activities during the TOEFL iBT® writing tasks: A stimulated recall study. *ETS Research Report Series*, *2015*(1), 1–42.

Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal*, *100*(1), 320–340.

Barkaoui, K. (2019). What can L2 writers' pausing behavior tell us about their L2 writing processes? *Studies in Second Language Acquisition*, *41*(3), 529–554.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Lawrence Erlbaum Associates, Inc.

Chapelle, C. A., Grabe, W., & Berns, M. (1993). *Communicative language proficiency: Definitions and implications for TOEFL 2000*. Educational Testing Service.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, *18*(1), 80–98. https://doi.org/10.1177/0741088301018001004

Choi, I., & Deane, P. (2021). Evaluating writing process features in an adult EFL writing assessment context: A keystroke logging study. *Language Assessment Quarterly*, *18*(2), 107–132.

Chukharev-Hudilainen, E. (2019). Empowering automated writing evaluation with keystroke logging. In E. Lindgren & K. Sullivan (Eds.), *Observing writing: Insights from keystroke logging and handwriting* (pp. 125–142). Brill.

Conijn, R., Roeser, J., & van Zaanen, M. (2019). Understanding the keystroke log: the effect of writing task on keystroke features. *Reading and Writing*, *32*(9), 2353–2374.

Cumming, A. (1989). Writing expertise and second-language proficiency. *Language Learning*, *39*(1), 81–135. https://doi.org/https://doi.org/10.1111/j.1467-1770.1989.tb00592.x

Cumming, A. (2016). Theoretical orientations to L2 writing. In M. M. Rosa & M. Paul Kei (Eds.), *Handbook of second and foreign language writing*. De Gruyter Mouton.

Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, *10*(1), 5–43.
https://doi.org/https://doi.org/10.1016/j.asw.2005.02.001

Deane, P. (2011). *Writing assessment and cognition (ETS Research Report RR-11-14)*. Educational Testing Service.

Deane, P., Odendahl, N., Quinlan, T., Fowles, M., Welsh, C., & Bivens-Tatum, J. (2008). Cognitive models of writing: Writing proficiency as a complex integrated skill. *ETS Research Report Series*(2), i–36.

Deane, P., Sabatini, J., Feng, G., Sparks, J., Song, Y., Fowles, M., . . . Foley, C. (2015). Key practices in the English language arts (ELA): Linking learning theory, assessment, and instruction. *ETS Research Report Series*.

di Gennaro, K. (2006). Second language writing ability: Towards a complete construct definition. *Working Papers in TESOL & Applied Linguistics*, *6*(2).

Ellis, R., & Yuan , F. (2004). The Effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, *26*, 59–84.
https://doi.org/10.1017/S0272263104026130

Flower, L., & Hayes, J. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, *31*, 21–32.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365–387.

Galbraith, D., & Baaijen, V. M. (2019). Aligning keystrokes with cognitive processes in writing. In (pp. 306–325). Brill. https://doi.org/https://doi.org/10.1163/9789004392526_015

Gebril, A., & Plakans, L. (2013). Toward a transparent construct of reading-to-write tasks: The interface between discourse features and proficiency. *Language Assessment Quarterly*, *10*(1), 9–27.

Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistics perspective*. Longman.

Guo, H., Zhang, M., Deane, P., & Bennett, R. (2020). Effects of scenario-based assessment on students' writing processes. *Journal of Educational Data Mining*, *12*(1), 19–45.

Hall, E. (1991). Variations in composing behaviours of academic ESL writers in test and non-test Situations. *TESL Canada Journal*, *8*, 9–33.

Hayes, J. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 6–44).

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, *29*(3), 369–388.
https://doi.org/10.1177/0741088312451260

Janssen, D., van Waes, L., & van den Bergh, H. (1996). Effects of thinking aloud on writing processes. In *The science of writing: Theories, methods, individual differences, and applications.* (pp. 233–250). Lawrence Erlbaum Associates, Inc.

Kellogg, R. (2001). Competition for working memory among writing processes. *The American journal of psychology*, *114*, 175–191. https://doi.org/10.2307/1423513

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57–71). Laurence Erlbaum Associates.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: An evaluation of the 1996 Model. *The Journal of Writing Research*, *5*, 159–190.

Khuder, B., & Harwood, N. (2015). Writing in test and non-test situations: Process and product. *Journal of Writing Research*, *6*(3), 233–278. https://doi.org/10.17239/jowr-2015.06.03.2

Krapels, A. (1990a). An overview of second language writing process research. In B. Kroll (Ed.), *Second language writing (Cambridge Applied Linguistics): Research insights for the classroom* (pp. 37–56). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9781139524551.007

Krapels, A. (1990b). The interaction of first and second language composing: Processes and rhetorics. *Dissertations Abstracts International*, *51*(12), 4045A.

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using inputlog to analyze writing processes. *Written Communication*, *30*(3), 358–392. https://doi.org/10.1177/0741088313491692

Manchón, R. M., Larios, J., & Murphy, L. (2007). A review of writing strategies: Focus on conceptualizations and impact of first language. *Language learner strategies: 30 years of research and practice*, 229–250.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3 ed., pp. 13–103). The American Council on Education/ Macmillan Publishing Co.

Michel, M., Révész, A., Lu, X., Kourtali, N.-E., Lee, M., & Borges, L. (2020). Investigating L2 writing processes across independent and integrated tasks: A mixed-methods study. *Second Language Research*, *36*(3), 307–334.

Miller, K. S. (2000). Academic writers on-line: investigating pausing in the production of text. *Language Teaching Research*, *4*(2), 123–148.

Miller, K. S., Lindgren, E. V. A., & Sullivan, K. P. H. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, *42*(3), 433–454. https://doi.org/https://doi.org/10.1002/j.1545-7249.2008.tb00140.x

Mislevy, R. J. (2007). Validity by design. *Educational Researcher*, *36*, 463–469.

Odendahl, N., & Deane, P. (2018). Assessing the writing process: A review of current practice. *ETS Research Memorandum Series*, *ETS-RM-18-10*.

Plakans, L. (2008). Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing*, *13*(2), 111–129.

Purpura, J. E. (2013). Cognition and language assessment. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1452–1476). John Wiley & Sons.

Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly*, *19*(2), 229–258.

Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language Learning*, *37*(3), 439–468. https://doi.org/10.1111/j.1467-1770.1987.tb00579.x

Roca de Larios, J., Manchón, R., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, *17*(1), 30–47. https://doi.org/https://doi.org/10.1016/j.jslw.2007.08.005

Roca de Larios, J., Marín, J., & Murphy, L. (2001). A temporal analysis of formulation processes in L1 and L2 writing. *Language Learning*, *51*(3), 497–538.

Roca de Larios, J., Murphy, L., & Manchon, R. (1999). The use of restructuring strategies in EFL writing: A study of Spanish learners of English as a foreign language. *Journal of Second Language Writing*, *8*(1), 13–44.

Roca de Larios, J., Murphy, L., & Marin, J. (2002). Critical examination of L2 writing process research. In S. Ransdell & M. L. Barbier (Eds.), *New Directions For Research in L2 Writing* (pp. 11–47). Cambridge University Press.

Roca de Larios, J., Nicolás-Conesa, F., & Coyle, Y. (2016). Focus on writers: Processes and strategies. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 267–286). De Gruyter Mouton. https://doi.org/doi:10.1515/9781614511335-015

Révész, A., Kourtali, N.-E., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, *67*(1), 208–241.

Sasaki, M. (2002). Building an empirically-based model of EFL learners' writing processes. In S. Ransdell & M. Barbier (Eds.), *New directions for research in L2 writing* (Vol. Studies in Writing: 11, pp. 101–122). Kluwer Academic Publishers.

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.

Silva, T., & Leki, I. (2004). Family matters: The influence of applied linguistics and composition studies on second language writing studies—Past, present, and future. *The Modern Language Journal*, *88*(1), 1–13.

Stratman, J. F., & Hamp-Lyons, L. (1994). Reactivity in concurrent think-aloud protocols: Issues for research. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology*. Sage Publications.

Weigle, S. C. (2002). *Assessing Writing*. Cambridge University Press.

Wolcott, W. (1987). Writing instruction and assessment: The need for interplay between process and product. *College Composition and Communication*, *38*, 40–46.

Worden, D. L. (2009). Finding process in product: Prewriting and revision in timed essay responses. *Assessing Writing*, *14*(3), 157–177.

Yang, H. C., & Plakans, L. (2012). Second language writers' strategy use and performance on an integrated reading-listening-writing task. *TESOL Quarterly*, *46*(1), 80–103.

Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarterly*, *16*(2), 195–209.

Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly*, *17*(2), 165–188.

Zhang, M., Zou, D., Wu, A. D., Deane, P., & Li, C. (2017). An investigation of writing processes employed in scenario-based assessment. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and investigating response processes in validation research* (pp. 321–339). Springer International Publishing. https://doi.org/10.1007/978-3-319-56129-5_17