Exploring task-based cognitive processes: Methodological advances and challenges

Andrea Révész, University College London

Abstract

This paper argues that TBLT researchers should dedicate more effort to investigating the cognitive processes in which L2 learners engage during task work to facilitate theory-construction and to inform pedagogical practices. To help achieve this, a review follows of various subjective (questionnaires, interviews, think-aloud/stimulated recall protocols) and objective (dual-task methodology, keystroke-logging, eye-tracking) methods that are available to TBLT researchers to examine cognitive processes underlying task-based performance. The paper concludes that, to obtain a more valid understanding of task-generated cognitive processes, it is best to combine various methods to overcome the limitations of each. Finally, some methodological recommendations are provided for future cognitively-oriented TBLT research.

Introduction

The last three decades have seen an increased interest in exploring the usefulness of tasks in promoting second language use and development. Most of this research has been concerned with investigating task-based products, assessing the quality of task performance in terms of general linguistic outcome measures such as complexity, accuracy, and fluency (CAF) and gauging task-based development with pretest-posttest measures of specific linguistic constructions. So far, much less research has focused on the cognitive processes in which L2 learners engage when they carry out tasks. Although some scholars had called for research into task-generated cognitive processes already in the 1990s (Mackey, 1999; Swain & Lapkin, 1995), research on task-based processes still remains sparse in the field of cognitively-oriented TBLT research. In this piece, I argue that, to facilitate TBLT theory-construction as well as pedagogical practices, it is crucial that researchers dedicate more attention to task-generated cognitive processes. Next, I review a variety of approaches, both more traditional and state-of-the art, to examining task-based processes, while highlighting the advantages and disadvantages of each method. Building on some recent work, I also describe and demonstrate how triangulating various methods can help us achieve fuller and more complete understanding of task-based processes and learning. Finally, I offer some recommendations for future TBLT research on cognitive processes.

Rationale for investigating task-based processes

The investigation of task-generated cognitive processes is important from both a theoretical and pedagogical perspective. On the theoretical front, examining task-based cognitive processes is essential to reach firm and valid conclusions about theoretical models of task-based performance and development. As Norris and Ortega (2003) point out and other validation frameworks (Kane, 2006; Messick, 1995) also highlight, to assess theoretical frameworks adequately, independent evidence needs to be provided for the validity of each and every construct included in them. This applies to any independent variables, causal processes, and dependent variables invoked by the models. Otherwise, it will not be possible to achieve valid and meaningful interpretations about them.

To give an example, the Limited Capacity Model (Skehan, 1998) and the Cognition Hypothesis (Robinson, 2001a), two widely studied task-based models of speech production and development, make several predictions about the effects of manipulating task complexity (independent variable) on CAF indices (dependent variable) as well as the cognitive processes mediating the presumed links between task complexity and CAF measures (causal processes). Therefore, if researchers would like to test the predictions of these frameworks, they need to provide specific and independent evidence for all the constructs entailed in them, starting with task complexity, continuing with the causal processes invoked, and ending with the CAF measures posited (Norris, 2010; Révész, 2014). In recent years, studies of task complexity increasingly include independent measures of cognitive load or mental effort into their design (e.g., Baralt, 2013; Lee, 2019; Malicka, 2020; Malicka & Levkina, 2012; Révész et al., 2014; Révész et al., , Michel, Gilabert, 2016; Sasayama, 2016; Torres, 2018; Zalbidea, 2017; Xu et al., in press), thereby providing validity evidence for the task complexity manipulations under scrutiny. A lot of progress has also been made to identify ways that may help increase validity when it comes to selecting CAF measures (Bulté & Housen, 2012; Housen & Kuiken, 2009; Housen, Kuiken & Vedder, 2012; Norris & Ortega, 2009). To date, however, relatively few studies have attempted to assess the causal cognitive processes, such as attentional allocation and speech production processes, postulated in the models. In the absence of such evidence, no valid conclusions can be reached about the predictions made by the Limited Capacity Model and the Cognition Hypothesis.

The same applies to other theoretical frameworks that TBLT researchers tend to draw on. For example, when task-based studies assess predictions made by Levelt's (1989) speech production model, direct evidence for processes such as conceptualisation, formulation, and monitoring should be provided. Similar, if researchers adopt Kellogg's (1996) model of writing as a theoretical basis when examining task effects on L2 writing, it is important to seek information about writing processes posited by Kellogg (e.g., planning, translation, and monitoring) to fully test any predictions made based on the model. In sum, to facilitate theory-building within TBLT, it is imperative that not only task-based outcomes, but also causal, task-generated cognitive processes are investigated.

In addition to advancing theory-building, information about task-based processes can also generate useful insights for second language instruction. For example, knowledge about links between observable task-based behaviours and associated cognitive processes may assist teachers in identifying sources of task difficulty, thus enabling them to fine-tune teaching to the specific needs of learners. For example, it is likely that a learner who is struggling primarily with planning what to say or write during task work will benefit from differential instruction than a learner who is struggling with linguistic encoding processes such as lexical and syntactic encoding. Research into task-based processes can also generate insights into the extent to which task-based work creates the type of learning opportunities that are assumed to promote L2 development.

Methods to investigate task-based processes

Having established the need to examine task-based processes, now we turn to a review of tools that TBLT researchers have available to look into cognitive processes generated during task-based work. There are two different types of methods that can be used to investigate cognitive processes: subjective approaches, which rely on learner self-reports of their thoughts and perceptions, and objective methods, which involve observations of learners' behaviours or brain activity during task performance. In what follows, I discuss various subjective and objective methods currently available to tap task-generated processes. In doing so, I will draw on my own and others' work to give representative examples how various methods can be employed to explore the cognitive processes involved in second language speaking, writing, reading, and listening task performance.

Questionnaires

Among the subjective techniques, the use of written self-report questionnaires has probably been the most widely used to examine task-based processes. Like in the larger field of second language acquisition (SLA) (Iwaniec, 2020), TBLT researchers have often elicited participants' perceptions about various aspects of their performance by the means of written questionnaires. The use of self-report questionnaires is based on the assumption that people can retrospectively report on the cognitive processes in which they have engaged during task performance.

Many task-based studies included closed-ended questionnaire items, which require respondents to select a response from a limited number of options (e.g., Baralt, 2013; Gilabert, 2007; Levkina & Gilabert, 2014; Michel, 2011; Robinson, 2001b). To illustrate, Révész (2011) used closed-ended items to obtain information about teachers' and students' perspectives on their subjective experience during simple and complex oral interactive tasks that students completed in a classroom environment. Among the questionnaire items, two concerned the attentional processes in which students engaged during task performance. In particular, both teachers and students were asked to assess on a 7-point scale which task version they found more effective in drawing the students' attention (1) to the quality of their output and (2) to the quality of their peers' production.

Fewer studies relied on open-ended questionnaire items to tap task-generated cognitive processes to get more detailed information about participants' experiences during task work. For example, Révész (2009) used open-ended questions to gain insight into participants' perspectives on task difficulty when performing tasks with or without visual support. In a more recent study, Sasayama and Norris (2019) also elicited open-ended responses, included in self-report measures, to assess participants' perceptions about task difficulty and mental effort during tasks which had been designed to generate differential cognitive demands. The participants were asked about what made the tasks they had performed easy/difficult and what made the tasks entail low/high mental effort (p. 100).

While questionnaire responses can yield interesting insights and have the advantage of being easy to administer, there are limitations associated with this method, especially if the researchers' aim is to elicit information about cognitive processes. A key issue is that participants, due to memory decay, might not have reliable memory of what they were thinking when they were completing the task. Also, questionnaire items often elicit short responses, which lack detail.

Subjective time estimation

Another subjective method that has been utilised in process-oriented TBLT research is subjective time estimation (e.g., Block et al., 2010, Zakay & Block, 1997). Subjective time estimation entails asking participants to estimate the amount of time they have spent carrying out a specific task without the use of an external timing device. According to a meta-analysis by Block et al. (2010) in the field of psychology, the interpretation of subjective time estimates depends on whether or not participants are made aware in advance that they will need to judge the duration of a task. If participants are asked to assess task duration retrospectively without being informed about the subsequent time estimation, a greater cognitive load is associated with an increase in the subjective-to-objective duration ratio, that is, the estimated time increases as a result of increasing processing demands. Simply put, when a person is working on a more cognitively demanding task, time is perceived to pass more slowly if they are asked to estimate time spent on task after task completion. On the other hand, if participants are made aware prior to task performance that they will need to estimate task duration, a greater cognitive load is related to a decrease in the subjective-to-objective duration ratio, that is, the estimated time elapsed will drop as cognitive demands increase.

This technique has been utilised in a number of studies in the field of TBLT (e.g., Baralt, 2013; Lee, 2019; Malicka & Levkina, 2012; Sasayama, 2016). Adopting the retrospective time estimation paradigm, Baralt (2013) asked her participants to provide a verbal estimation of how long a particular version of a task took after task completion. Malicka and Levkina (2012) also requested their participants to give retrospective time judgments. In addition, participants were asked to estimate which of the two task versions under investigation took them longer to complete. Sasayama (2016) opted for the use of prospective time estimation. Given that participants were asked to estimate time spent on task in the practice phase of the experiment, they were likely to expect that they would need to give time estimations during the actual experiment. The analytical procedures in these three studies, however, are not directly comparable to those employed in studies considered in Block et al.'s (2010) meta-analysis. As Lee (2019) pointed out, in cognitive psychology subjective-to-objective time ratios are typically computed, whereas these L2 studies compared raw difference scores calculated by subtracting the actual time from the guess time participants provided (guess time - actual time). Lee addressed this issue by using subjective-to-objective time ratios when studying native speaker task performance using the prospective time estimation approach.

Interviews

Interviews, another subjective technique, can also be employed to gain insights into learners' cognitive processes during task work. This technique has been used in some SLA (Sachs & Polio, 2007) and TBLT studies, but it has been less popular than verbal protocols such as the think-aloud and stimulated recall procedure (see below) to gain information about learners' cognitive processes. Among the various interview types, TBLT researchers have predominantly relied on semi-structured interviews when investigating task-generated cognitive processes. While addressing a prepared set of questions, this interview format enables the researcher to ask follow-up questions and thereby let the participant elaborate on issues of interest.

A study by Ortega (2005) provides a good example of how interviews can be used to gain in-depth information about participants' cognitive processes. Ortega employed semi-structured interviews to elicit information from L2 learners of Spanish about their strategic planning processes prior to task performance. As typical of semi-structured interviews, Ortega prepared a protocol to guide the interview process but also probed deeper into participants' thoughts by the means of follow-up questions. These probes were often inspired by the researchers' observations and field notes taken in the pre-task planning phase.

A more recent study by Pang and Skehan (2014) also illustrates well how interviews can be utilised to investigate participants' cognitive processes during pre-task planning. The researchers interviewed L2 learners of English about how they used a 10-minute planning period to prepare for their subsequent narrative performance. During the interviews, the participants were first prompted by some general questions about what they did while they planned. These were followed by more focused prompts, which aimed to tap what specific areas participants paid attention to during the planning stage, whether their emphasis was on words, grammar, ideas, structure of the story, rehearsal, or clarity of expression.

Both of these studies yielded detailed and in-depth insights into participants' thoughts during pre-task planning. Indeed, interviews are more prone to elicit elaborate and detailed responses than questionnaires. As compared to questionnaires, however, interviews have the disadvantage of being more time-consuming to set up, conduct, and analyse. In addition, the limitation remains that participants might not fully remember during the interview what they were thinking when they actually engaged in task work.

Think-aloud protocols

To decrease the risk of memory decay, a few TBLT researchers have relied on introspective methods such as the think-aloud procedure to shed light on task-based cognitive operations.

Think-aloud protocols, as the name of the procedure suggests, involve participants in thinking aloud concurrently as they perform an activity; they are asked to verbalise whatever comes into their mind while they are carrying out the task (Bowles, 2010). This subjective technique has been used less often by TBLT researchers than researchers in the broader field of instructed SLA. This might be partly due to the fact that most TBLT research has so far focused on speaking tasks, and think-aloud protocols, as they themselves involve speaking, cannot be employed to investigate cognitive processes underlying listening and speaking task performances.

Nevertheless, some TBLT research has demonstrated that think-alouds can be successfully used to gain insights into participants' thought processes concurrently during task-based work. Sangarun (2005), another study on pre-task planning, was among the first to utilise the think-aloud procedure to investigate task effects. The aim of this study was to examine how strategic planning processes may differ as a function of planning instruction. The participants were Thai L2 English participants, who received one of three different instruction types to plan: to focus on meaning, focus on form, or focus on meaning and form. The experiment started with a 15-minute training session on how to think aloud. Then, participants had 15 minutes to plan each of their task performance, a monologic instruction and argumentative task. The think-aloud data allowed Sangarun to obtain detailed information about participants' planning processes under the various planning conditions.

A clear advantage of using think-alouds in this study was that there was no threat of memory loss. In general, during think-alouds there is little danger of inaccurate reporting due to memory decay given that participants are simultaneously reporting what they are thinking. One potential disadvantage of think-alouds, however, is that they may result in reactivity. This means that the act of thinking aloud may change the cognitive processes in which learners engage and thereby risk construct validity (see, however, Bowles, 2010; Leow et al., 2010). Another potential issue associated with think-aloud protocols is veridicality; verbal reports may not be able to capture all of the participant's thought processes.

Stimulated recall protocols

Like think-aloud protocols, the stimulated recall procedure is another introspective technique that can be used to investigate task-generated cognitive processes. This method intends to elicit the thoughts that participants had while performing a task through *a posteriori* recall sessions (Gass & Mackey, 2017). Researchers use some reminder (e.g., video-recording of task performance) to prompt participants' recall.

Although L2 researchers have used this technique extensively in several areas of instructed SLA research to investigate topics such as attention/awareness/noticing, corrective feedback perception, and strategy use (see Gass & Mackey, 2017 for a review), few studies have utilized the stimulated recall technique to examine the effects of task-related variables (e.g., Révész, Kourtali et al., 2017; Torres, 2018). Among the few exceptions is Kim et al.'s (2015) study, which investigated how task complexity influences the cognitive processes in which L2 learners engage during oral task-based interaction. Learners' oral interactions were video-recorded during simple and complex tasks, and immediately after, they were invited to describe the thoughts they had during task performance. To stimulate recall, the researcher conducting the recall session replayed the video-recording of participants' performance and encouraged them to stop the recording whenever they wanted to comment on their behaviours. In addition, the researcher stopped the recording when the learners seem to have experienced difficulty or displayed an interesting behaviour when engaged in the original task performance.

Clearly, in Kim et al.'s research, the verbal protocol did not interfere with participants' task performance. As compared to think-alouds, the stimulated recall procedure carries less risk in terms of reactivity. Stimulated recalls, however, pose more substantial threat to veridicality. As a result of memory loss, participants are less likely to give entirely accurate and full reports of the cognitive processes in which they had previously engaged. Nevertheless, if the technique is carefully implemented, potential issues of veridicality can be considerably mitigated (Gass & Mackey, 2017).

Dual-task methodology

Given the limitations associated with subjective methods, more recently TBLT researchers have also begun to use objective methods to tap task-based processes. Dual-task methodology is one objective method that TBLT researchers started to employ to tap cognitive processes during task performance. Dual-task methodology entails performing a secondary task concurrently with the primary task. In the context of TBLT, primary tasks take the form of pedagogic tasks, and secondary tasks typically include simple activities that demand sustained attention, such as detecting simple auditory (Brünken et al., 2004) or visual (Cierniak et al., 2009) stimuli. The principle underlying the technique is that the amount of cognitive load generated by the primary task is mirrored in the accuracy and speed with which the secondary task is carried out. Slower or less accurate secondary task performance is taken to reflect greater cognitive demands posed by the primary task.

While dual-task methodology had been used by SLA researchers for other purposes in the past (e.g., Declerck & Kormos, 2012; DeKeyser, 1997), it was not until relatively recently that this technique began to be utilised in TBLT research. Studies by Révész et al. (2014), Révész et al. (2016), and Sasayama (2016) were among the first to employ dual-task methodology to investigate cognitive task demands during task performance. In Révész et al.'s (2014) work, the primary tasks asked participants to select one of two past events in a famous person's life and then produce a past counterfactual statement orally about the event. The events were displayed on the computer screen with the help of key phrases and pictures. Under the low-complexity condition, the choice between the two events was designed to be simple, requiring little reasoning. During the complex version of the task, however, the choice between the two events was intended to be more complex, posing greater reasoning demands. Révész et al. (2016) included low- and high-complexity versions of three types of primary tasks (a decision-making, a map, and a narrative task), with each being presented to the participants via computer. In both studies, during the secondary task the colour of the background screen randomly changed to green or red, and the participants were asked to respond as fast and accurately as possible when the colour changed to green while disregarding changes to red. In Sasayama's (2016) work, the primary tasks involved performing four computer-delivered narrative tasks based on picture stories. The tasks were constructed to involve differential cognitive demands through manipulating the number of elements included in the narratives. The secondary task asked participants to react as fast and accurately as possible when a change occurred in the colour of a capital letter (*A*), which appeared above the picture story on the participants' screen. Participants had to respond to changes from black to red but not to react to changes from red to black. In each study, it was expected that the task versions intended to be more cognitively demanding would lead to slower and less accurate secondary task performance.

All three studies, along with some more recent work (e.g., Lee, 2019; Xu et al., in press), found that dual-task methodology, if operationalised and implemented carefully, can be useful for identifying differences between actual cognitive demands and intended mental effort or difficulty. Dual-task methodology is also regarded as a reliable and sensitive methodology in the area of cognitive psychology; it has the advantage of providing a concurrent and objective measure of processing load. One disadvantage of the technique, however, is that it is relatively obtrusive. Therefore, studies employing dual-task methodology need to include a baseline condition, which enables researchers to observe how participants would perform on the primary task in the absence of a secondary task.

Keystroke-logging

Keystroke-logging is another objective tool that TBLT researchers can rely on when studying L2 writing processes. Keystroke-logging programs register all the keystrokes and mouse movements that writers produce. The resulting log files then can be used for further analyses to obtain detailed information about concurrent writing behaviours, for example, by the means of fluency, pausing, and revision indices (Lindgren & Sullivan, 2019; Van Waes, Leijten, Lindgren, & Wengelin, 2016).

Keystroke logging was already used more than 20 years ago to study the impact of task-related variables on L2 writing behaviours. In a pioneering study, Spelman Miller (2000) employed keystroke-logging software to examine whether fluency and L2 pausing behaviours were different across evaluative and descriptive essay writing tasks. While the participants, L2 writers of English and L1 English writers, produced the two essays, their keystrokes were recorded. Then, the researcher analysed the log data for several fluency and pausing measures. Another seminal keystroke-logging study by Thorson (2000) examined how the revision behaviours of L2 writers vary by task genre. All the participants wrote a newspaper article as well as a letter in L1 English and L2 German, during which their keystrokes were captured. More recently, researchers have started to use keystroke-logging more widely to investigate L2 task effects. For example, Révész, Kourtali et al. (2017) have employed keystroke logging, in combination with stimulated recall, to explore the impact of a task complexity manipulation on the speed fluency, pausing, and revision behaviours of L2 writers. Keystroke-logging has also been utilised to look into how writing behaviours may vary across independent and integrated writing tasks, so far mostly focusing on testing tasks (Barkaoui, 2016; Michel et al., 2020). Another area where keystroke-logging is beginning to be used is the study of text-chat interactions. Charoenchaikorn's (2019) study provides a good example of how keystroke-logging can be utilised in this context. The researcher assessed participants' revision behaviours, speed fluency, and accuracy during task-based text-chat interactions based on logs obtained through keystroke-logging software.

It is not surprising that task-based researchers show an increased interest in using keystroke-logging to study writing processes. Being an unobtrusive data collection tool and generating real-time data, keystroke logging has several benefits in comparison to verbal protocols that have traditionally been used to investigate cognitive writing processes. Unlike verbal protocols, however, keystroke-logging data have the drawback of supplying no direct information about L2 writers' cognitive activities. Additionally, data gathered via keystroke logs cannot reveal insights about the reading processes in which L2 writers engage.

Eye-tracking

Eye-tracking constitutes another objective tool that can help us make inferences about task-generated cognitive processes, especially those that occur during tasks involving interaction with some type of visual input such as a reading text, picture prompt, or video (Conklin et al., 2018; Godfroid, 2019). Eye-tracking entails recording the participant's moment-to-moment movement of their gaze or eye fixations while they carry out a task. The principle underlying eye tracking is that where, in what order, and how long the participant fixates their eyes mirrors the attentional processes in which they engage when processing visual input (Reichle, 2006).

For the past decade, a growing number of TBLT studies have employed eye-tracking to investigate attentional allocation during task-based work, mirroring a trend in the larger field of instructed SLA. One area of task-related eye-tracking research focused on exploring the effects of task complexity manipulations on how learners distribute their attention during task performance. For example, Révész et al.'s (2014) previously discussed study employed eye-tracking, besides dual-task methodology, to investigate whether, as predicted, the complex version of the task would lead to quantitatively and/or qualitatively different eye-movements as compared to the simple version of the same task. The researchers employed two types of measures frequently used in eye-tracking: number of eye-fixations (i.e., how many times participants fixated at an area of interest) and duration of eye-fixations (i.e., how long the fixations were). It was hypothesized that, in the complex version, participants would spend more time looking at the area including the two alternative pictures (see above for a description of the task manipulation), given that a more complex decision had to be made.

In another, less controlled study of task complexity, Michel et al. (2014) examined how learners interacted with picture prompts during simple versus complex versions of three types of oral tasks, as part of the same project on which Révész et al.'s (2016) earlier described study was based. The researchers expected that, under the complex task conditions, participants would display longer and more eye-fixations, as they would need to work out more complex relationships (decision-making task), take into account more elements in more intricate arrangements (map task), or engage in more intentional reasoning (narrative task). Vasylets and Gilabert (2015), using data elicited by the same decision-making task as Michel et al. (2014) and Révész et al. (2016), conducted more sophisticated eye-tracking analyses to elucidate the potential effects of task complexity. Instead of relying on the whole picture prompt as an area of interest, the researchers defined their interest areas according to what factors were likely to differentiate participants' attentional processes during simple and complex task performance. To give an example, the decision-making task asked participants to explain what actions they would take in case of a fire in a building. Differences between the low- and high-complexity versions included whether people in the building were part of vulnerable groups (e.g., pregnant/elderly people) and how much equipment was available (e.g., fire engines, helicopters). Therefore, the interest areas comprised the people in the building, the fire engines, and helicopters; and the researchers compared fixation durations and counts for these areas under the simple and complex task conditions.

Focusing on oral task performance as well, McDonough et al. (2015) utilised eye tracking to explore another aspect of attentional allocation, whether joint attention achieved through eye gaze predicts L2 speakers' responses to corrective feedback during task-based interaction. The researchers coded for how long the feedback-provider gazed at the L2 speaker when giving feedback, and how long the L2 speaker gazed at the feedback-provider during their response to the feedback received. The analyses also involved coding for whether there was mutual eye gaze between the speaker and the interlocutor providing feedback. In the next step, researchers explored the relationship between these eye-movement measures and learners' responses to feedback, whether they provided no response or produced a target-like or a non-target-like reformulation. McDonough et al. (2020) replicated this research to address some limitations of the original study such as the relatively small sample size and issues with the experimental set-up (better placement of webcams).

A few studies have also used eye-tracking methodology to investigate features of task-based interaction in text-chat environments. Smith (2012) was the first to explore the potential of eye-tracking in examining the noticing of corrective feedback by L2 learners during text-chat interaction tasks. The researcher measured noticing in terms of stimulated recall comments and increased visual attention, which was operationalised as learners' total fixation time on recasts provided by a native speaker. Michel and Smith (2019) employed eye-tracking to study alignment, one's tendency to re-use their interlocutor's language patterns, during task-based text-chat interaction. The researchers first identified all examples of multi-word units (three- to ten-word *n-grams*) that both interlocutors produced. In case participants fixated on an overlapping n-gram at least on one occasion, this n-gram was identified as an area of interest. The eye-tracking analyses involved comparing participants' fixation times and counts for these interest areas with those for turns that their interlocutors had produced in the same text-chat interaction but resulted in no overlap. In a further text-chat study comparing alignment during tutor and peer task-based interaction, Michel and O'Rourke (2019) followed similar procedures in analysing the eye-tracking data collected. Notably, however, the authors in this study combined eye-tracking with cued interviews to gain fuller insights into attentional processes.

Another area within TBLT where researchers have begun to use eye-tracking is the study of task-generated L2 writing processes. Michel et al.'s (2020) previously mentioned study, for example, employed stimulated recall protocols and keystroke-logging together with eye-tracking to capture differences in composing processes across independent and integrated TOEFL iBT writing tasks. The researchers defined their area of interest as the whole writing window in the TOEFL iBT environment, as the spacing did not enable carrying out word-level eye-tracking analyses (see, however, Révész, Michel et al., 2017; Révész et al., 2019). Given the large interest area, the study has yielded relatively coarse-grained eye-tracking results. Nevertheless, the eye-movement data were helpful in obtaining some insights into differences in eye-movements during the two different task types.

Finally, it is worthwhile to highlight how eye-tracking may be exploited to explore attentional processes during multi-modal input-based tasks, where learners receive information through various channels such as audio, pictorial, and written input. In a recent study, Lee and Révész (2020) operationalised multimodal input-based tasks with the help of captions. The treatment task asked the participants to play the role of a newsroom editor, who was responsible for classifying news items. Participants had first watched a news clip, then their task was to evaluate whether a given category and title were appropriate for the news item. For one of the experimental groups, the target construction, the present perfect, was enhanced in the subtitles, whereas for the other experimental group there was no enhancement. To examine whether, as hypothesised, the enhanced constructions in the captions would attract greater amount of attention, the researchers calculated a variety of eye-tracking indices (number of visits, first pass reading time, second pass reading time, and skipping rate; see Conklin et al., 2018; Godfroid, 2019 for a review and definition of these eye-tracking measures).

In sum, eye-tracking appears to be a useful tool to investigate attentional processes in task-based work where visual stimuli play a key role. Advantages of this method include that it is an unobtrusive, concurrent, and online technique. An apparent limitation of eye-tracking methodology, however, is that it provides no direct insights into learners' cognitive processes. In other words, eye-gaze recordings can only give information about what learners are looking at and for how long, they cannot reveal why learners engage in the viewing behaviours observed. Another drawback to the use of eye-tracking is that sophisticated eye-tracking equipment is still not easily accessible for many.

Neuroimaging

Although neuroimaging has not yet been utilised in TBLT research, it is worth considering its potential to investigate task-generated neural processes based on research in other areas of L2 performance and learning. Functional magnetic resonance imaging or functional MRI (fMRI) appears to be one technique that might have the capacity to generate useful insights about the neural processes in which learners are involved during task-based work. fMRI scans detect increases in blood flow that accompany enhanced brain activity. When more blood is delivered to a part of the brain, this is reflected in scans captured by the fMRI machine.

Previously, L2 researchers have primarily used fMRI to examine L2 users' performance on highly controlled activities such as priming tasks (Pliatsikas et al., 2014). There are a few studies, however, which suggest that this method might also be suitable for assessing neural processes during task-based oral production. Jeong et al. (2016), for example, examined the neural processes involved in producing communicative versus descriptive utterances to identify brain areas relevant to communicative versus non-communicative speech production. Participants watched short videos, in which a person was handling an object or using a tool (e.g., playing the guitar). Under the communicative condition, participants were asked to casually talk to the actor in the video as if they were talking to him in daily life (e.g., asking "what kind of music are you playing"?). In the descriptive condition, on the other hand, they were instructed to describe the actor's situation (e.g., "he is playing the guitar"). The researchers also included a baseline condition in which the participants simply named the object in the video to control for articulatory motor processing. In the fMRI analyses, brain activation patterns were compared across the communicative and descriptive conditions. Building on Jeong et al.'s and others' work (e.g., Sassa et al., 2007), fMRI scans might also be utilised in TBLT research, for instance, to detect speech production processes such as conceptualisation and formulation. This is so because certain areas of the brain can be linked to intending to speak and planning a speech act (left posterior SMG), whereas other areas have been shown to be involved in linguistic processing (left middle frontal gyrus, dorsal part of the left inferior cortex, and opercular part of the left inferior frontal gyrus) (Jeong et al., 2016).

In light of this, it would seem that neuroimaging techniques might be useful for revealing task-induced brain activation patterns. The fMRI technique, in particular, has the benefit of yielding direct and objective insights into neural processes. The data obtained, however, fare low in terms of ecological validity; carrying out tasks in an fMRI scanner is quite different from completing tasks in classroom or real-life contexts. Also, for now, access to fMRI equipment is costly.

Triangulation of sources

Having reviewed some key methods to study task-generated cognitive processes, it is obvious that each and every one of them has advantages and disadvantages. To overcome the limitations associated with different techniques, TBLT researchers increasingly gather and triangulate various data sources, parallel with a generally growing trend to combine methods in instructed SLA research (King & Mackey, 2016; Gass & Mackey, 2017). To study the cognitive processes in which learners engage during task-based oral production, researchers have used together methods such as self-perception questionnaires, stimulated recall, dual-task methodology, and/or eye-tracking (e.g., Lee, 2019; Michel et al., 2014; Révész et al., 2014; Révész et al., 2016; Sasayama, 2016; Sasayama & Norris, 2019). Scholars working on text-chat task-based interaction have triangulated data gathered through interviews, the stimulated recall procedure, keystroke-logging, and/or eye-tracking (e.g., Charoenchaikorn, 2019; Michel & O'Rourke, 2019). Similar, studies investigating cognitive processes during task-based writing have combined insights from stimulated recall protocols, keystroke-logging, and eye-tracking (e.g., Michel et al., 2020; Révész, Kourtali et al., 2017; Révész, Michel et al., 2017; Révész et al., 2019; Stiefenhöfer & Michel, 2020). Bringing together multiple data-collection instruments, given the weaknesses inherent in each elicitation tool, helped produce, in each of these studies, a more comprehensive and valid picture of the cognitive processes underlying L2 task performance than the use of a single tool would have allowed for. Now a description and discussion follows of two pieces of TBLT research to illustrate how triangulation of sources can assist in achieving a more thorough account of task-based processes.

To date, probably the most comprehensive study in terms of triangulation has been the project presented in Michel et al.'s. (2014) and Révész et al.'s (2016) previously mentioned work. The researchers triangulated data gathered by the means of four methods – dual-task methodology, eye-tracking, self-perception questionnaires, and stimulated recall – to assess the impact of task complexity manipulations on task-generated cognitive load and processes. The results from both objective and subjective measures included in the study, overall, converged on the finding that the versions of the tasks intended to be more cognitively demanding indeed required greater cognitive effort. On each of the three task types (decision-making, map, and narrative), the participants displayed lower accuracy on the secondary task while completing the high-complexity versions of the primary tasks, rated the high-complexity tasks as requiring more mental effort, and provided stimulated recall comments indicating greater effort at conceptualisation during the task versions designed to be more cognitively complex. The eye-tracking indices also found some evidence for greater attention to the visual prompts on the complex decision-making and map tasks. Importantly, however, for the narrative task, secondary reaction times, eye-tracking indices, and task difficulty self-ratings yielded no significant difference as a function of the task complexity manipulations. In sum, although evidence was accumulated for the validity of the task complexity manipulation for all three tasks, the project generated more convincing validity evidence for a difference between the high- and low- complexity versions of the decision-making and map tasks. From a methodological perspective, it is important to highlight that, if the researchers had only relied on eye-tracking or task difficulty self-ratings, they would have failed to capture some evidence for a difference between the simple and complex narratives and, as a result, would have arrived at the interpretation that the two task versions did not differ in terms of cognitive load. Vice versa, if only secondary accuracy rates, mental effort-ratings, and stimulated recall comments had been included in the design, the validity evidence for the narrative task-complexity manipulation would have emerged stronger than justified, in light of the overall picture that had emerged when all data sources had been considered.

A study by Révész et al. (2019) also demonstrates how triangulation of data elicitation techniques can be used to investigate task-generated cognitive processes. One aim of this research was to examine L2 writers' pausing behaviours and associated cognitive writing processes when completing an IELTS essay writing task. It was hypothesised that pauses between lower textual units (e.g., pauses within and between words) would be more often related to lower-level, linguistic encoding processes (e.g., lexical retrieval and morphological encoding), whereas pauses between higher textual units (e.g., between sentences) would more often correspond to higher-order writing processes (e.g., planning content and organisation). To test these predictions, the study employed three data elicitation tools: keystroke logging, eye-tracking, and stimulated recall protocols. By the means of the keystroke-logging data, the researchers were able to assess how long participants paused and where they paused, whether the pauses occurred within words, between words, or between sentences. Keystroke logging alone, however, would have supplied no information about the writers' viewing behaviours during pauses, resulting in relatively limited insights into writing processes, as rereading previously written text is a principal writing mechanism involved in idea generation and monitoring. The use of eye-tracking enabled the authors to deal with this possible limitation, allowing for an inspection of where participants were looking when they paused. Specifically, participants' eye movements were coded according to whether their eye gaze(s) stayed during the pause at the point of inscription (i.e., the leading edge of the text written so far) or visited the word/phrase, clause, sentence, or paragraph prior to the point of inscription. Finally, the stimulated recall comments, unlike keystroke-logging and eye tracking, had the capacity to offer a window into the conscious thought processes of the participating L2 writers. As part of the stimulated recall protocols, participants, among other things, were asked to share what they were thinking when they paused. As expected, the study found that, when writers paused between sentences, it was more probable that they engaged in planning content and looked back on longer texts. In contrast, during pauses within and between words, writers tended to engage in linguistic encoding processes and view shorter texts such as rereading a word or phrase. Unlike previous research, however, they confirmed these patterns from complimentary sources using a single dataset, which afforded more valid inferences about the cognitive processes associated with pausing behaviours than a single method would have made possible.

Recommendations for further research

Although some TBLT researchers have already used a variety of methods and combined these in innovative ways to examine cognitive processes during task work, it is worthwhile to consider how the methodological aspects of future research in this area could be further advanced.

First, TBLT research on cognitive processes would benefit from more sophisticated triangulation of data sources. In studies of task-based processes so far, researchers have usually obtained separate, group-level summaries from the various data collection methods and have triangulated the results at the group level. In the future, researchers could also attempt to triangulate data at the individual level to achieve more detailed insights about task-based processes, using a combination of subjective and objective measures to overcome the limitations of various techniques.

Second, future studies could also conduct more detailed, time-locked analyses of the process data rather than drawing conclusions based on summaries of process data. For example, researchers could calculate eye-tracking and keystroke-logging indices for every minute or even every second of the task as opposed for the whole duration of task performance. In this way, information could be obtained about task-based processing on a moment-to-moment basis to capture the dynamic nature of task performance (e.g., Bygate, 2013). A few TBLT researchers have already taken steps in this direction. For example, Baralt and Gurzynski-Weiss (2011), in a study of state anxiety during task-based interaction, measured learners' anxiety half-way and right after task-based interaction self-report state anxiety questionnaires. Similar, researchers in the area of L2 writing (e.g., Roca de Larios et al., 2008) often compare processes at different stages of the writing process by typically dividing it into 3-5 stages.

Third, like in other areas of instructed SLA research, more studies are needed to investigate longitudinal changes in task-generated cognitive processes. In the past two decades there has been an increasing number of studies examining task-based development, but the majority of these studies were still short-term, and only few have included process-based measures (e.g., Lee & Révész, 2020). To inform TBLT theory and guide task-based teaching, however, it is also important to conduct longer-term studies on task-generated cognitive processes. Without doubt, carrying out longitudinal studies is challenging due to various practical constraints.

Finally, to further TBLT research in general and on task-based processes in particular, it is crucial that TBLT researchers continue to share their instruments in open-science platforms (e.g., IRIS) to facilitate the transparency and replication of TBLT studies.

References

Baralt, M. (2013). The impact of cognitive complexity on feedback efficacy during online versus face-to-face interactive tasks. *Studies in Second Language Acquisition, 35*, 689–725.

Baralt, M., & Gurzynski-Weiss, L. (2011). Comparing learners' state anxiety during task-based interaction in computer-mediated and face-to-face communication. *Language Teaching Research*, *15*, 201–229.
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. *Modern Language Journal*, *100,* 320–340
Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta- analytic review. *Acta Psychologica, 134,* 330–343.
Bowles, M. (2010). *The think-aloud controversy in language acquisition research*. Routledge.
Brünken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science, 32*, 115–132.
Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21– 46). John Benjamins.
Bygate, M. (2013, October). *On fetters and goals, and the development of an empirical TBLT in terms of language, learning and teaching*. Plenary speech presented at the 5th Biennial International Conference on Task-based Language Teaching, Banff, AB, Canada.
Charoenchaikorn, V. (2019). *L2 revision and post-task anticipation during text-based synchronous computer-mediated communication (SCMC) tasks.* Unpublished doctoral dissertation, Lancaster University, UK.
Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior, 25*, 315–324.
Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge University Press.
Declerck, M., & Kormos, J. (2012). The effect of dual task demands and proficiency on second language speech production. *Bilingualism: Language and Cognition*, 15, 782–796.
DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morpho-syntax. *Studies in Second Language Acquisition*, 19, 195–221.
Gass, S. M., & Mackey, A. (2017). *Stimulated recall methodology in applied linguistics and L2 research*. Routledge.
Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching, 45*, 215–240.
Godfroid, A. (2019). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. Routledge.
Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, *30*, 461–473.
Housen, A., F. Kuiken, & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken and I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1–20). John Benjamins.
Iwaniec, J. (2020). Questionnaires: implications for effective implementation. In J. McKinley, & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 324–335). Routledge.

Jeong, H., Sugiura, M., Suzuki, W., Sassa, Y., Hashizume, H., & Kawashima, R. (2016). Neural correlates of second-language communication and the effect of language anxiety. *Neuropsychologia*, *84*, 2–12.

Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). American Council on Education.

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. 57–72). Lawrence Erlbaum Associates.

Kim, Y., Payant, C., & Pearson, P. (2015). The intersection of task-based interaction, task complexity, and working memory: L2 question development through recasts in a laboratory setting. *Studies in Second Language Acquisition, 37,* 549–581.

King, K., & Mackey, A. (2016). Research methodology in second language studies: Trends, concerns and new directions. *The Modern Language Journal, 100(s),* 209–227.

Lee, J. (2019*).* Task complexity, cognitive load, and L1 speech. *Applied Linguistics*, *40*, 506–539.

Lee, M., & Révész, A. (2020). Promoting grammatical development through captions and textual enhancement in multimodal input-based tasks. *Studies in Second Language Acquisition, 42*, 625–651.

Leow, R., Grey, S., Marijuan, S., & Moorman, C. (2014). Concurrent data elicitation procedures, processes, and the early stages of L2 learning: A critical overview. *Second Language Research, 30*, 111–127.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

Levkina, M., & Gilabert, R. (2014). Task sequencing in the L2 development of spatial expressions. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 37–70). Bloomsbury.

Lindgren, E., & Sullivan, K. (Eds.). (2019). *Observing writing: Insights from keystroke logging and handwriting* (Vol. 38). Brill.

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21, 557–587.

Malicka, A. (2020). The role of task sequencing in fluency, accuracy, and complexity: Investigating the SSARC model of pedagogic task sequencing. *Language Teaching Research*, *24*, 642–665.

Malicka, A. & Levkina, M. (2012). *Measuring task complexity: does L2 proficiency matter?* In A. Shehadeh, & C. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and implementation* (pp. 43–66). John Benjamins.

McDonough, K., Crowther, D., Kielstra, P. & Trofimovich, P. (2015). Exploring the potential role of eye gaze in eliciting English L2 speakers' responses to recasts. *Second Language Research, 31*, 563–575.

McDonough, K., Trofimovich, P., Dao, P., & Abashidze, D. (2020). Eye gaze and L2 speakers' responses to recasts: A systematic replication study of McDonough, Crowther, Kielstra and Trofimovich (2015). *Language Teaching*, *53*, 81–95.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749.

Michel, M. (2011). Effects of task complexity and interaction on L2 performance. In P. Robinson (Ed.), *Second language task complexity: researching the Cognition Hypothesis of language learning and performance* (pp. 141–174). John Benjamins.

Michel, M., & O'Rourke, B. (2019). What drives alignment during text chat with a peer vs. a tutor? Insights from cued interviews and eye-tracking. *System*, *83*, 50–63.

Michel, M., Révész, A., & Gilabert, R. (2014). *Eye movement prompts in stimulated recall: tapping cognitive processes based on audio vs. visual stimuli.* Paper presented at AILA, Brisbane, Australia.

Michel, M., Révész, A., Lu, X., Kourtali, N.-E., & Borges, L. (2020). Investigating L2 writing processes across independent and integrated tasks: A mixed-methods study. *Second Language Research, 36*, 243–255.

Michel, M., & Smith, B. (2019). Measuring lexical alignment during L2 chat interaction: An eye-tracking study. In S. Gass, P. Spinner, & J. Behney (2019) (Eds.), *Salience in second language acquisition* (pp. 244–268). Routledge.

Norris, J. M. (2010, September). *Understanding instructed SLA: Constructs, contexts, and consequences.* Plenary address delivered at the annual conference of the European Second Language Association (EUROSLA), Reggio Emilia, Italy.

Norris, J., & Ortega, L. (2003). Defining and measuring SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 717–761). Blackwell.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, *30*, 555–578.

Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 77–109). John Benjamins.

Pang, F., & Skehan, P. (2014). Self-reported planning behaviour and second language performance in narrative retelling. In Skehan, P. (Ed.), *Processing perspectives on task performance* (pp. 95–128). John Benjamins.

Pliatsikas, C., Johnstone, T., & Marinis, T. (2014). fMRI Evidence for the Involvement of the Procedural Memory System in Morphological Processing of a Second Language. *PLoS ONE*, *9*(5), e97298.

Reichle, E. D. (2006). Theories of the “eye-mind” link: Computational models of eye-movement control during reading. *Cognitive Systems Research, 7*, 2–3.

Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition, 31*, 437–470.

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *Modern Language Journal*, *95*, 162–181.

Révész, A. (2014). Towards a fuller assessment of cognitive models of task-based learning: Investigating task-generated cognitive demands and processes. *Applied Linguistics*, *35,* 87–92.

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

Révész, A., Michel, M., & Gilabert, R. (2016). Measuring cognitive task demands using dual task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition*, *38*, 703–737.

Révész, A, Michel, M., & Lee, M. (2017). Investigating IELTS Academic Writing Task 2: Relationships between cognitive writing processes, text quality, and working memory. *IELTS Research Reports Online Series*, 2017/3.

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behaviours: A mixed methods study. *Studies in Second Language Acquisition, 41*, 605–631.

Révész, A., Sachs, R., & Hama, M. (2014). The effects of task complexity and input frequency on the acquisition of the past counterfactual construction through recasts. *Language Learning, 64*, 615–650.

Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for investigating task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22,* 27–57.

Roca de Larios, J., Manchón, R. M., 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,* 30–47.

Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 111–141). John Benjamins.

Sasayama, S. (2016*)*. Is a 'complex' task really complex? Validating the assumption of cognitive task complexity. *The Modern Language Journal, 100*, 231–254.

Sasayama, S., & Norris, J. (2019). Unravelling cognitive task complexity: Learning from learners' perspectives on task characteristics and second language performance. In Z. Wen, & J. A. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honour of Peter Skehan* (pp. 95–132). John Benjamins.

Sassa, Y., Sugiura, M., Jeong, H., Horie, K., Sato, S., & Kawashima, R. (2007). Cortical mechanism of communicative speech production. *NeuroImage*, *37,* 985–992.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Smith, B. (2012). Eye tracking as a measure of noticing: A study of explicit recasts in SCMC. *Language Learning & Technology, 16*, 53-81.

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

Stiefenhöfer, L.**,** & Michel, M. (2020). Investigating the relationship between peer interaction and writing processes in computer-supported collaborative L2 writing: A mixed-methods study. In R. M. Manchón (Ed.), *Writing and language learning: Advancing research agendas* (pp. 255–280). John Benjamins.

Swain, M. & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*, 371–391.

Thorson, H. (2000). Using the computer to compare foreign- and native-language writing processes: A statistical and case study approach. *The Modern Language Journal, 84*, 55–70.

Torres, J. (2018). The effects of task complexity on heritage and L2 Spanish development. *Canadian Modern Language Review, 74*, 128–152.

Van Waes, L., Leijten, M., Lindgren, E., & Wengelin, Å. (2016). Keystroke logging in writing research: Analyzing online writing processes. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 410–426). The Guilford Press.

Vasylets, L., & Gilabert, R. (2015). *Exploring the visual-dynamic and linguistic conceptualisation traces in task-based performance*. Paper presented at the 6th Biennial International Conference on Task-based Language Teaching, Leuven, Belgium.

Xu, T. S., Zhang, L. J., & Gaffney, J. S. (in press). Examining the relative effectiveness of task complexity and cognitive demands on students' writing in a second language. *Studies in Second Language Acquisition*.

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science*, *16*, 12–16.

Zalbidea, J. (2017). One task fits all? The roles of task complexity, modality, and working memory capacity in L2 performance. *The Modern Language Journal*, *101*, 335–352.