Stimulated recall

Hugo Santiago Sanchez and Trevor Grimshaw

Introduction

Stimulated recall was first used as a method of data elicitation in the 1950s and has been employed ever since across a wide range of disciplines, including nursing, counselling, medicine, psychotherapy, social psychology and education. It is variously referred to as stimulated recall, stimulus or stimulated recall interviews, or stimulated recall analysis. In this chapter we will use the term ‘stimulated recall’ throughout.

Stimulated recall is an introspective research method which has most often been used to serve one of two main purposes: either the recall and elicitation of interactive thought processes (Gass & Mackey, 2017), or ‘the recall of events to facilitate a discussion of the factors influencing them’ (Borg, 2006, p. 219). To assist these recall processes, a stimulus is used, such as video or audio recordings. Typically, participants watch or listen to video/audio recordings and then provide verbal commentaries on them, either spontaneously or prompted by the researcher.

The emergence of new fields of inquiry and the diversification of research aims, approaches, contexts, resources and issues in recent years have motivated new understandings and uses of stimulated recall to meet specific local needs, characteristics and challenges. The purpose of this chapter is to offer a critical review of the literature on stimulated recall, focusing specifically on empirical studies published recently, with a view to exploring current conceptualisations of this method in applied linguistics research. This includes the diverse methodological purposes that stimulated recall serves, the multifaceted and multimodal data collection procedures it involves, and the epistemological and practical challenges it presents.

Key literature on stimulated recall

Within the extensive literature on stimulated recall there are certain publications that are cited with particular frequency. As such, they may be regarded as seminal or relatively authoritative, and might constitute, we believe, a good starting point for applied linguists who are contemplating this method of research.
Gass and Mackey (2017) offer a detailed description and critique of stimulated recall in the general context of second language research. They begin by contextualising stimulated recall within the broader range of introspective methods. They then go on to provide guidance on how to design a study based on the use of stimulated recall to elicit concurrent thoughts, how to analyse the data collected and how stimulated recall may be combined with other data collection methods. They conclude with a discussion of the limitations of stimulated recall and some recommendations as to how to offset these. A useful accompaniment is Ryan and Gass (2012), which is based on a case study of how stimulated recall was applied to explore miscommunication between L1 and L2 users of English in a university setting. Ryan’s account of the study is followed by Gass’s commentary on the emergent problems. Together, the two authors offer valuable insights into key issues such as validity and reactivity.

Borg (2006) provides a summary of several studies that have employed stimulated recall for eliciting teachers’ verbal commentaries, and critically examines some of the key issues involved in this kind of research. Borg himself has made extensive use of this method for gathering retrospective accounts of teachers’ cognitions, thus gaining insights into the mental processes that inform their decision-making. Borg adds a different perspective on the purpose of this research method to that of Gass and Mackey (2017), namely to facilitate a discussion of the factors which shape teachers’ classroom decisions and actions.

Three empirically based publications shed further light on key issues in the application of stimulated recall in second language learning and teaching research. Egi (2008) reports on the use of this method for the elicitation of L2 learners’ cognitive processes and performance in oral interaction and collaborative learning activities. The author focuses specifically on the causes of reactivity, which may include the nature of stimuli, the impact of verbalisation on the learners’ performance and the combination of these factors. Polio, Gass & Chapin (2006) report on the use of stimulated recall as a means of exploring native-English speakers’ perceptions of non-native speaker interactions. The paper includes a discussion of the limitations of stimulated recall and proposes some refinements. Modiba and Stewart (2014) report on a study of professional development in South Africa in which classroom observation was combined with stimulated recall discussion. The authors argue that this combination of methods was effective in encouraging an English first additional language teacher to ‘take stock’ and reflect critically on his teaching. The study focused specifically on the extent to which the teacher’s practice aligned with current educational policy.

Four other publications fall outside the field of applied linguistics but enjoy frequent citation by language researchers. Drawing on the broader tradition of qualitative social research, Messmer (2015) argues for the value of stimulated recall as a method for investigating teachers’ thinking processes and actions. Rowe (2009) reports on a study within the field of music education which involved the use of video-stimulated recall to explore gender expectations. This includes a brief but useful reflection on using the method with child participants. A very different set of issues is discussed by Dempsey (2010) in his vivid account of how stimulated recall was used in ethnographic research on jazz musicians’ jamming. The author comments on the interactive nature of stimulated recall interviews and the fact that participants were often surprised by elements of the recordings that were played back to them. Finally, Lyle (2003), in his account of research on non-deliberative decision-making by sports coaches, offers a detailed evaluation of the benefits and limitations of stimulated recall research.

Although these publications were the products of diverse contexts and disciplines, their accounts of procedures and their discussions of the challenges and issues in stimulated recall are readily applicable to the field of applied linguistics.
A content analysis of recent studies using stimulated recall

To explore current conceptualisations of stimulated recall in applied linguistics research, we focused on recent high-impact research published in English which utilised stimulated recall as a method for collecting data. In identifying relevant journals, we drew on the 2016 Impact Report published by Thomson Reuters and, within this, we applied the filters ‘Social Science Citation Index’ and ‘Linguistics’. This produced a list of 182 journals in the fields of theoretical and applied linguistics, which we then reviewed individually (from highest to lowest journal impact factor), through a close examination of the aims and scope described in their homepages, so as to identify the top 30 highest ranked journals in applied linguistics.

The subsequent phase involved the identification of relevant articles from these top 30 highest ranked journals. Our searches used a range of keywords (stimulated/stimulus recall [interview], recall [tool/method], memory retrieval) and focused on issues published between January 2012 and April 2018. A total of 196 publications were identified and then analysed to select original research articles about studies that had adopted stimulated recall and to exclude irrelevant items (e.g. editorials, literature reviews, book reviews, disciplinary dialogues and forum sections). This resulted in a total of 88 papers selected for data extraction and analysis.

We acknowledge that in limiting our review to journal articles published in English we may have missed important empirical studies, such as those published in book chapters and non-English language journals as well as unpublished research (e.g. doctoral dissertations), which could potentially showcase original rationales, approaches and designs for employing stimulated recall for data collection. We feel confident, however, that our review was based on a substantial body of primary research which collectively provides insights into the latest rationales and practices concerning the use of stimulated recall in applied linguistics research.

Data were extracted from each article and placed in a systematic map for analysis which included broad themes, such as purpose, participants, stimulus, procedure, strengths and limitations. The data were then subjected to content analysis procedures that involved codification, thematic analysis and categorisation of themes.

Discussion of themes identified in recent stimulated recall research

Table 26.1 shows the thematic categories and subcategories which emerged from our content analysis. Each constitutes a potential aspect or issue which an applied linguist might need to consider when designing and implementing stimulated recall research. In the following theme-based sections we will present the most salient findings and cite example studies from our sample throughout.

Contexts, disciplinary fields and participants

The geographical contexts in which the studies were conducted varied greatly, covering six continents and over 20 countries. Within these contexts, most of the studies were in the areas of second/foreign/additional language education. Other fields included teacher education (Yuan & Lee, 2014), manga writing (Robertson, 2017) and editorial work (Wiley & Tanimoto, 2013).

Similarly, the studies involved a wide range of participant types. Most commonly, the participants were educators or students. The educators were mostly in-service language teachers at various levels, from preschool (Palviainen & Mård-Miettinen, 2015) to higher education (Stickler & Shi, 2017). Other studies involved pre-service teachers (Yuan & Lee, 2014), a writing teacher (Watson, 2015) and university tutors (Heron, 2018).
Table 26.1 Themes and subcategories emerging from the analysis

<table>
<thead>
<tr>
<th>SR Themes</th>
<th>Subcategories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td></td>
</tr>
<tr>
<td>Discipline</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>age group</td>
</tr>
<tr>
<td></td>
<td>roles</td>
</tr>
<tr>
<td></td>
<td>number</td>
</tr>
<tr>
<td></td>
<td>language proficiency</td>
</tr>
<tr>
<td>Purpose</td>
<td>types</td>
</tr>
<tr>
<td>Stimulus</td>
<td>episodes/events</td>
</tr>
<tr>
<td>Interview</td>
<td>procedures</td>
</tr>
<tr>
<td>procedures</td>
<td>interview type</td>
</tr>
<tr>
<td></td>
<td>method</td>
</tr>
<tr>
<td></td>
<td>mode</td>
</tr>
<tr>
<td></td>
<td>means</td>
</tr>
<tr>
<td></td>
<td>language choice</td>
</tr>
<tr>
<td></td>
<td>time</td>
</tr>
<tr>
<td></td>
<td>duration</td>
</tr>
<tr>
<td>Training</td>
<td>reported use</td>
</tr>
<tr>
<td></td>
<td>purpose</td>
</tr>
<tr>
<td></td>
<td>amount</td>
</tr>
<tr>
<td></td>
<td>type</td>
</tr>
<tr>
<td></td>
<td>critiques</td>
</tr>
<tr>
<td>Research issues</td>
<td>credibility</td>
</tr>
<tr>
<td></td>
<td>procedures</td>
</tr>
<tr>
<td></td>
<td>strengths</td>
</tr>
<tr>
<td>Practical</td>
<td></td>
</tr>
<tr>
<td>implications</td>
<td></td>
</tr>
</tbody>
</table>

In studies with learner participants, the majority were undergraduates or postgraduates (Negretti & McGrath, 2018). In a few cases the participants were adolescents at secondary school (Rutgers & Evans, 2017). Most student participants were learners of English as a second, foreign or additional language, while some were learners of other languages or bilingual learners (Rutgers & Evans, 2017).

Other participants included learning advisors (McCarthy, 2018), English-teaching editors at Japanese universities (Willey & Tanimoto, 2013), multilingual writers (Kobayashi & Rinnert, 2013), air traffic controllers interacting in a virtual testing environment (Park, 2018), bilingual families following the one-parent-one-language method (Venables, Eisenclhas, & Schalley, 2014) and the author of Japanese manga (Robertson, 2017).

None of the studies in our sample involved child participants. This might be explained, we would argue, by the limited attentional strategies, processing capacity and memory store of children, especially young ones (Pinter, 2011).

The numbers of participants involved in the studies tended to be relatively small: often no more than ten people. This is, of course, characteristic of qualitative research. In a few studies the numbers were larger, ranging from 30 to 77. The participants often constituted a sample that was drawn from a larger number of participants and where the study involved multiple methods of data collection.
Purposes of stimulated recall research

As stated in our introduction, the literature identifies stimulated recall as serving one of two major purposes. The first of these is the use of stimulated recall to elicit interactive or concurrent cognitions, particularly ‘thoughts’ (Gass & Mackey, 2017). In line with this, many of the studies in our sample focused on the elicitation of thoughts when participating in some interaction (Nakatsukasa, 2016) or taking some action (Gurzynski-Weiss & Baralt, 2014). The lead question guiding these stimulated recall interviews is: What were you thinking about right then?

However, in recent years researchers have used stimulated recall in connection with an increasingly broader and more sophisticated range of cognitive or affective constructs and processes. These include ‘thinking processes and intentions’ (Hsieh, 2017), ‘thoughts and awareness’ (Jackson & Cho, 2018), ‘thought processes and motivations’ (Palviainen & Mård-Miettinen, 2015), ‘inner thoughts and affective responses’ (Qiu & Lo, 2017) and ‘cognitive and metacognitive strategies’ (Park, 2018).

The second major purpose of stimulated recall involves the recall of events as a springboard to discussing the factors that influenced those events. Within the seminal literature on stimulated recall opinions on this matter are sharply divided. Gass and Mackey (2017) believe that one of the risks of stimulated recall interviews is that participants may wander ‘off track’ and end up reflecting on broader issues. In contrast, Borg (2006) argues that this is precisely where stimulated recall interviews are most useful – as the basis for reflective discussion.

Here, again, the articles in our sample exhibited much variation. Several studies focused on the interpretations and rationales that teachers offered for their pedagogical decisions, including the reasons that underpinned their use of new technologies in class (Van Praag & Sanchez, 2015), grammar teaching techniques (Sanchez & Borg, 2014) and bilingual language practices (Palviainen & Mård-Miettinen, 2015).

Other studies sought explanations or interpretations of students’ learning behaviours, including the use of genre knowledge in writing (Negretti & McGrath, 2018) and the provision of peer feedback (Yu & Lee, 2014, 2016). The affective dimension of language learning was addressed in studies such as Gregersen, Meza, and MacIntyre (2014), which elicited students’ explanations for fluctuations in their level of language learning anxiety. Some researchers elicited the problems that learners had encountered when performing a specific task and explored their responses to those difficulties (Fernandez Dobao, 2012).

Some studies involved the use of stimulated recall with both teachers and students, with the purpose of comparing their perspectives. This was the case, for example, of Stickler and Shi (2013), which investigated learner and teacher strategies.

In addition to these two commonly cited purposes of stimulated recall, this method has been used, we would argue, for at least two further purposes. One of them involves the combination of the two preceding purposes: to elicit interactive cognitions and to facilitate discussion. Examples of studies with this combined purpose include Heron (2018), which elicited university tutors’ thoughts at the time of an action, invited them to reflect on interactions and explored their values and attitudes; and Yu and Lee (2014), which sought to elicit university students’ thoughts while engaging in peer feedback tasks, the beliefs that underpinned their experiences and their rationales.

A further purpose of stimulated recall is to promote critical self-reflection. Examples from the studies we analysed include Tavares (2015), in which a teacher was prompted to reflect on the extent to which learning outcomes had been met; Venables, Eisenchlas, & Schalley (2014), in which bilingual families were encouraged to reflect on parent–child interactions and their motivations for the use of particular languages; and Johnson (2015), in which educators were invited to comment on critical moments in team-taught lessons.
Stimulated recall

Three main categories of stimulus can be identified in the sample of articles. The first of these involves technology-based artefacts, among which video recording is the one cited most widely. Typically, participants are asked to watch either entire video recordings or extracts of these.

Recent developments in technology have enabled the use of new forms of video recording. These include split-screen video recordings, whereby, for example, the left-hand side of the video screen displays the student perspective and the right-hand side shows the teacher’s perspective. Jackson and Cho (2018, p. 35) state that ‘[s]imultaneous viewing[s] . . . may serve as stronger stimuli for the reminder of events and prompt participants’ recall through the provision of situational and contextual support’.

Another important innovation has been the use of gaze-plot video recordings (i.e. eye tracking – see Pellicer-Sánchez & Conklin, this volume, for full discussion). Through the accumulation of information, it is possible to create a record of key focal points or ‘heat areas’ on the screen. As Stickler and Shi (2017, p. 172) argue, this ‘can add detail to our understanding of online behaviour, thereby making reflections potentially more powerful and more meaningful’. Audio recordings have also been used extensively in stimulated recall studies. Other technology-based stimuli include online chat transcripts or chat logs and computer screen recordings.

The second category of stimulus consists of materials produced or used by the participants. Most of these are developed naturally in the research context, such as study logs, learning notebooks, essays, articles, assignments, edited abstracts and tutor or peer written feedback. Others are adopted by the participants as part of their tasks, such as course materials and tests.

The third category comprises stimuli that are specifically produced by the researcher for research purposes. These include lesson transcripts, field notes, questionnaire responses, self-reported anxiety ratings and bitmap graphs.

Gass and Mackey (2017) argue that stimuli can be strengthened by using more than one source: for example, by watching a video in conjunction with reading a transcript. Several researchers in our sample followed this approach of combining stimuli. This included combinations of, for example, audio recordings and transcripts (Heron, 2018); edited abstracts, transcripts and audio recordings (Willey & Tanimoto, 2013); and field notes, transcripts and video recordings (Gierlinger, 2015).

Stimuli are used to support the recall of specific ‘critical’ episodes or events which bear immediate relevance to the focus of the study. In our sample, most of these incidents were teaching-related and included lesson excerpts, teaching demonstrations, oral tasks, collaborative class activities, writing processes or behaviours, reading tasks, classroom or virtual interactions, interactions within the family, corrective feedback, peer feedback, learning advisory sessions, mentoring meetings, online tutorials and simulations.

The incidents were selected in a variety of ways. In most studies the researcher chose the events either prior to or during the interview. This was sometimes complemented with the researcher and the participants choosing additional incidents during the interview (Kim, 2014) or with the participants pausing the recording when they identified an event they wanted to comment on (Sanchez & Borg, 2014). In some cases the events were chosen by the participants themselves during the course of the interview, for example, by stopping the audio every time a critical event occurred (Fernandez Dobao, 2012). In relatively few studies the events were selected collaboratively by the participants and the researcher. This was done to ‘[avoid] the risk of appearing judgemental or selecting unflattering extracts’ (Whyte, Schmid, van Hazebrouck Thompson, & Oberhofer, 2014, p. 135).
Though the procedures for selecting and using stimuli and events are still clearly dominated by the researcher, there is a growing tendency to appreciate and enable participants’ agency and empowerment. This, we believe, contributes to strengthening the quality of the data collected and the ethical procedures followed.

**Interview types and procedures**

A range of different stimulated recall interview types and procedures were reported in the articles. Where authors reported on the type of interview used in their studies, these tended to be open-ended interviews (free verbalisations in reaction to the video recording). None of the studies reported the use of structured interviews.

Researchers variously reported the use of open questions (Heron, 2018), prompt questions (Qiu & Lo, 2017), probing and confirmation questions (Lai, Hu, & Lyu, 2018) or questions that explicitly encouraged reflection (Johnson, 2015). Others explained that they had adopted a fairly passive role, asking only clarification questions (Palviainen & Mård-Miettinen, 2015).

The most common method was to interview participants individually (Jackson & Cho, 2018). Other variations involved interviewing the participants in pairs (Hsieh, 2017), groups (Johnson, 2015), or a combination of individuals and pairs (Fernandez Dobao, 2012).

Some researchers described the preparation for the interview meetings. In some cases, for example, the participants and researchers viewed the video independently before the meeting (Whyte et al., 2014). In other studies lesson transcriptions were given to the participants to review before the stimulated recall interview took place (Watson, 2015).

In all but three cases the mode of communication for interviews was oral and the means was face-to-face. The exceptions were one study that employed email interviews (Robertson, 2017), one that used a combination of oral and written interviews (Tomita & Spada, 2013) and one that involved a combination of face-to-face and telephone interviews (Willey & Tanimoto, 2013).

The choice of language used for recall is a crucial factor. Gass and Mackey (2017) point out that when the language of the original event is different from the language of recall, this constitutes a particular challenge. As they put it, the key question is: ‘What is the potential influence of the language used in the recall sessions on the accuracy of the recall?’ (p. 47). The problem may be even more acute if the participants have limited proficiency in the target language.

Most of the researchers in our sample conducted their studies in the participants’ L1 (Fernandez Dobao, 2012). As some researchers explained, this was to ensure that the participants could avoid linguistic difficulties and could fully express their opinions (Qiu & Lo, 2017). In two articles the researchers report using the participants’ L2. One of them acknowledges that, ‘[b]ecause of the relatively low L2 proficiency of the participants, the learners might not have been fully able to express their thoughts’ (Nakatsukasa, 2016, p. 792).

Several researchers reported that the stimulated recall was conducted in the participants’ L1 and/or L2. Some stated that this was the choice of the participants (De Silva & Graham, 2015). One group of researchers gave the stimulated recall instructions in both the L1 and the L2, then allowed the participants to respond in whichever language felt most natural (Barkaoui, Brooks, Swain, & Lapkin, 2013). Gass and Mackey (2017) acknowledge the benefits of this approach but issue the caveat that the participants may have varying levels of proficiency, and so the data may not be of comparable quality.

The time lapse between the original event and the recall session is another factor that is of crucial significance in stimulated recall research. Gass and Mackey (2017) recommend that the recall session should be held as soon as possible after the original event. In our sample
Stimulated recall

many researchers conducted this immediately after the event, particularly when the stimulated recall was based on tasks performed by the participants (Hsieh, 2017). This approach has certain limitations. For example, Gierlinger (2015), who conducted the stimulated recall after lessons, was forced to rely on his field notes and memory of lesson episodes. In the majority of the articles in our sample, the recall sessions were held within seven days of the initial event (e.g. Kobayashi & Rinnert, 2013), though in a few studies they were conducted more than a week or two afterwards (e.g. Heron, 2018).

Of those studies that mentioned the duration of interviews, the majority reported that these took between 30 and 60 minutes (Sanchez & Borg, 2014). Some interviews were reported as being shorter, while a few lasted over an hour.

Training

In nine articles authors report that they had provided their participants with some form of training or guidance. The type of training varied between instructions in participants’ L1 and L2 (Barkaoui et al., 2013), video demonstrations of how to recall thoughts (Jackson & Cho, 2018) and a warm-up task (Rassaei, 2013).

Authors explained that the purpose of the training was to ensure participants’ full comprehension (Barkaoui et al., 2013), to help them to verbalise their cognitive/metacognitive processes during the performance of a task (Park, 2018) or to provide practice in reporting thoughts while watching a video (Rassaei, 2013).

Arguably, the training of participants may have an impact on the quality of data gathered by stimulated recall (Borg, 2006). However, only one of the articles in the sample suggests possible risks, warning that training may overly guide recall and lead participants to produce certain ‘desired’ responses (De Silva & Graham, 2015).

Research issues

Stimulated recall research does not come without its limitations. Karimi and Norouzi (2017, p. 46) are amongst several writers who acknowledge the ‘inherent subjectivity in SR technique’. The credibility of data has been questioned on account of the possibility of post hoc rationalisations (Watson, 2015). Reflecting on their own study, De Silva and Graham (2015, p. 49) acknowledge that the process of recall ‘may lead the respondent to “imagine” strategy use that did not in fact take place’.

Another limitation is that the participants may simply forget what they were thinking at the time when they were engaging in the event, or that their memory may at least deteriorate. Thus, McCarthy (2018, p. 184) comments that ‘[i]t was not possible to guarantee the accuracy of advisor verbalizations due to factors such as being selective in relating experiences or lapses in memory’. Likewise, Gurzynski-Weiss and Baralt (2014, p. 32) acknowledge that participants might ‘experience memory deterioration once they have finished a task’ and that ‘the accessibility of learners’ mental processes is, admittedly, not an undisputed issue’.

Several authors in our sample identified reactivity as a key issue. McCarthy (2018, p. 184) for example, notes the ‘possibility that participants may have (knowingly or unknowingly) altered their perception of reality for purposes of self-aggrandizement’. Qiu and Lo (2017, p. 695), in turn, acknowledge that ‘the participants might have altered their second performances or their reports due to their awareness of being observed. They might guess the potential assumptions and tend to provide more positive responses to the variables’.
Gass and Mackey (2017, p. 50) recommend that, to increase the accuracy of responses, the recall questions should ‘focus on the time frame in question’. They suggest the following as appropriate questions:

- What were you thinking when she said $x$?
- What were you thinking when you shook your head?
- I notice that you shifted your position in your chair and you hesitated. What were you thinking about then?

Similarly, Barkaoui et al. (2013, p. 311) warn against asking hypothetical questions, stating that ‘[the participants were] reminded that they should report what they were thinking at the time, not what they thought they should have thought or done, or how they thought they should have responded’.

In addition, certain practical constraints may impact on stimulated recall procedures. One of these relates to the question of who selects the episodes that are to be commented upon. Whyte et al. (2014), for example, found that delegating this choice to teachers was not always possible because: (a) practitioners varied in their definitions of the term ‘best practice’; (b) some teachers were afraid of displaying technical, pedagogical or organisational challenges; (c) some lacked confidence; and (d) copyright violations could occur during the recorded lesson, thus limiting the choice of extracts.

Researchers may also find themselves limited in terms of the type of episode that can be used. For instance, in their study of the thought processes of a Japanese multilingual writer, Kobayashi and Rinnert (2013) state that:

> due to time constraints and to lighten the burden of her recalling what she did during shorter pauses, we decided to take up only pauses 5 seconds or longer, which made it easier for her to explain more accurately what she was thinking at the time.

**Concluding comments and recommendations**

This review set out to explore current conceptualisations of stimulated recall in applied linguistics research. The data show that stimulated recall has developed in recent years beyond its original limited applications and enjoys more variety and flexibility in its design and uses, with some new practices following a more participant-oriented approach. These qualities have made stimulated recall appropriate for use in a wide range of geographical contexts and disciplines, and with a variety of participant types.

A notable development in stimulated recall research involves the expansion of its aims beyond the elicitation of interactive thoughts. More recent purposes of stimulated recall encourage participants to engage in critical reflection and discussion. This has had implications not only for research but also for participants’ personal and professional development. For instance, Hussin (2013, p. 112), reflecting on her study of simulated patient interactions, notes ‘the value that the students [gain] from the actual research process itself’. Likewise, Stickler and Shi (2017) highlight the importance of stimulated recall for raising students’ awareness of learning strategies, thus helping them to reflect upon their approaches and to expand their performative repertoires.

Technological advances have also impacted on how stimulated recall research is designed and applied. Technology-based stimuli have become more sophisticated, for example, with
split-screen video recordings enabling simultaneous viewings of teacher and student perspectives (Jackson & Cho, 2018), or eye tracking techniques creating a record of focal points and facilitating the investigation of online behaviour (Stickler & Shi, 2017). This sophistication of the stimuli and of the recall process has therefore opened up new perspectives in applied linguistics research and made the collection of introspective data more effective. In addition, technology has contributed to the stimulated recall interview process, for example, by adding multimodality (e.g. allowing interviews to be conducted online) and flexibility, thus expanding the reach of research contexts and participants. In view of all this, we believe that greater consideration should be given to the opportunities for innovative research that are afforded by new technologies, for example, mobile technologies.

In addition, there is an overall trend for recent stimulated recall interview procedures to be more participant-oriented in nature. For example, though the selection of critical episodes is still often made by the researcher, there is evidence of many participants being invited to choose the events themselves (either prior to or spontaneously during the interview) or to collaborate with the researcher on identifying relevant and meaningful incidents. Similarly, some studies report giving participants the choice of the language to be used for recall. These practices promote, we would argue, participant agency, empowerment and satisfaction, and add to the quality of research by enhancing data richness and ethical procedures. However, we believe that the choice of language(s) used in stimulated recall studies is worthy of further consideration. Most studies report the use of a monolingual approach to stimulated recall interviewing (using either the participants’ L1 or L2), with only a handful facilitating a cross-lingual or multilingual approach. This practice does not sit well with the heteroglossic nature of many research settings or the recognition of translanguaging as an interactional norm. Moreover, the current movement towards researching multilingually suggests important implications in terms of not only the quality of data collected but also the ethics of offering participants with opportunities for expression.

In relation to the epistemological and practical challenges of stimulated recall research reported by some of the studies reviewed, these are not new and have been well documented in the literature for decades, along with suggestions on how to address them (see, for example, Gass & Mackey, 2017). Despite its limitations, stimulated recall is now well established as one of the major approaches to introspective investigation. As McCarthy (2018, p. 184) notes, ‘[stimulated recall] has been validated in the research literature as a reliable introspection tool, which allows for claims to be made about participants in real-life settings’.

References


