Michael Yeldham* and Rainbow Tsai-Hung Chen

Investigating mediation styles of second language listener verbal reports

Abstract: This study investigated verbal report styles for eliciting strategies data from second language listeners. It examined outcomes from three different mediation (prompting) styles, one style unprompted and the two others prompted, after the learners were first provided with low-prescriptive instructions on how to complete the report. Also, the unprompted style was additionally examined after the provision of more-prescriptive instructions to observe the effect of this greater learner guidance. Theoretically, the core of the study examined two competing cognitive perspectives on verbal reporting. One, from an information processing perspective, is that verbal reports elicit the best insight into individuals' strategic processes when prompts are kept to a minimum. The other perspective, a constructivist one, advocates the use of prompts in the form of researcher questions, mainly to help guide the report. Seventeen Taiwanese EFL learners participated in the study, with data gathered from each through a verbal report followed by a semi-structured interview. It was found that researcher prompting was both strongly favored by the learners and clearly elicited the best data for second language listener strategies research. The results also indicated that unprompted reports were little more effective when preceded by more-prescriptive instructions.

Keywords: second language listening, verbal report, verbal report mediation, listening strategies

1 Introduction

Verbal report (VR) is widely viewed as a direct way of examining the cognitive processes learners use when performing a task, to the extent that it is now a staple of L2 research (Bowles 2010; Cohen 2011). However, a continuing issue in its use concerns the degree to which respondents ought to be prompted to elicit their thoughts (Cohen 2011; Ericsson and Simon 1993; Fox et al. 2011). While undergoing many language use tasks learners are able to concurrently report their thoughts

*Corresponding author: Michael Yeldham, Centre for Applied English Studies, Rm 6.47, Run Run Shaw Tower, The University of Hong Kong, Pokfulam, Hong Kong,
E-mail: mayeldham@hotmail.com
Rainbow Tsai-Hung Chen, Foreign Language Center at National Chengchi University, Taiwan

as think-alouds (Bowles 2010). However, this is not possible while listening, so listeners must deliver their VRs retrospectively to the task – preferably immediately after listening to the relevant utterance, when their thoughts are most accessible from working memory (WM) (Gruba 1999; Macaro et al. 2007). Given this need for retrospective reporting, the issue of prompting (mediation) is particularly relevant to listener VRs, as researcher questioning may be required to help focus the listener's report on the area of interest to the study and to stimulate the listener's retrieval of his or her thoughts (Jourdenais 2001). Consequently, this study examines the effects of different forms of mediation used to elicit L2 learners' thoughts, specifically their use of strategies during listening tasks.

2 Mediation in verbal reports

This research examines the issue of mediation in listener VRs from a cognitive perspective, where there are two main schools of thought. The first of these, an information processing view, was buoyed following extensive examination of VRs by Ericsson and Simon (1984, 1993). Their investigation focused on VRs used mainly in problem-solving tasks from cognitive science, where the reporting of the exact sequence of one's cognitive processes is considered important. One conclusion from their investigations was that for both concurrent and retrospective protocols, researcher prompts in the form of questions tend to interfere with the respondent's VR. Essentially, they argued that listening to or answering questions may deflect the respondents' attention from the thoughts they heeded during the task, thus leading to the loss of some thoughts from memory or distortion of the thoughts by other thoughts subsequently heeded after the task. Consequently adherents to this information processing view (Ericsson and Simon 1993; Fox et al. 2011), including applied linguists such as Green (1998), advise researchers to limit prompts to those encouraging the individual to keep talking.

To compensate for this absence of questions these adherents suggest clarifying the reporting task requirements for the respondent through pre-VR directions, such as "Say out loud everything in your mind as you solve the problem." Such adherents also stress that respondents only be asked to verbalize the thoughts they had during the task, and not be asked to explain the underlying reasons for them, which Fox, et al. (2011: 319) claim may "change both the sequence of thoughts and, by inference, measured performance." In the absence of researcher questions, verbalization warm-up tasks are also seen as important, to accustom learners to expressing themselves. Research by Kim (2002) suggests this is paramount for learners of Asian origin, who Kim found were more reluctant to express their thoughts than learners of European origin.

The second main cognitive view, a constructivist one, is based on the belief that comprehending a text is primarily a constructive endeavor (Afflerbach and Cho 2009; Cohen 2011; Gruba 1999; Pressley and Afflerbach 1995; Vandergrift 1998a). From this perspective, prompting through researcher questioning during the VR is seen as acceptable, and advocated mainly to help guide learners on what to report. Constructivists Pressley and Afflerbach (1995) first highlighted the discrepancy between their view and Ericsson and Simon's (1993) information processing view, by pointing to the difference between the "pure" problem solving tasks focused on by Ericsson and Simon, and the "fuzzy" and "multidimensional" nature of text comprehension (Pressley and Afflerbach 1995: 127). Indeed, Yeldham and Chen (2014) claimed such multidimensional factors in a listening task to include: the person's interpretation of the text (its gist and details); the strategies used to comprehend the text and the specific tactics employed to operationalize them (Goh 2002); whether the strategies are used before, during or after listening to the text (Vandergrift 2003) and whether they relate to wide sections of the text or more localized parts of it; and the person's thoughts about the text content.

Consequently, constructivists argue that relying entirely on pre-sessional instructions to guide the VR, as Ericsson and Simon (1993) suggest, may be sufficient to focus a VR on problem solving tasks such as mathematics problems, but not on complex, multi-dimensional tasks like text comprehension. Specifically, many such adherents argue that relying entirely on pre-sessional instructions may lead to respondents reporting information irrelevant to the research or that is too vague or generalized, or to omit reporting processes that *are* valuable to the research (Cohen 1998, 2011; Jourdenais 2001). Cohen (1998), and Pressley and Afflerbach (1995) also add that compensating for this deficiency through the use of highly prescriptive pre-sessional instructions may serve to move the learner away from naturalistic cognition and bias processing. Consequently, many constructivists contend that the use of mediation may help to avoid many of these pitfalls, and for listeners, who must deliver their VR retrospectively to the task, the addition of carefully considered questions may be especially needed to help them recall their thoughts (Jourdenais 2001).

3 Verbal report methods in past L2 listening studies

Over the past few decades, research has focused on examining two main aspects of listeners' cognitive processes. One has been the strategies listeners use to help

them construct meaning, while the other has looked at various aspects of listeners' bottom-up processing, particularly the skills involved in decoding connected speech. Strategies are processes which are generally conscious to the learners, while bottom-up skills are processed more automatically, so strategies are thus more suitable for research through VR (Buck 1990; Cohen 2011; Færch and Kasper 1987). As a result, the vast majority of past L2 listening studies featuring the use of VRs have centered on the learners' strategy use. A number of these studies have adopted Ericsson and Simon's (1993) information processing approach, where researcher questioning during the VR has been absent or minimal, with the respondent guided instead by pre-VR instructions, warm-up tasks and/or a preliminary training session (Goh 1998, 2002; Graham 1997, Graham et al. 2008, 2011; Laviosa 1991; Murphy 1985; Young 1997).[1] Note that the warm-up tasks were used primarily to accustom respondents to verbalizing their thoughts, and included listening tasks, mathematics problems and verbal imagery tasks.

By contrast, many of the other studies of listener strategies have followed the constructivists' path, and used questions, sometimes quite liberally, to help prompt learner response (while also using pre-VR instructions, often using warm-up tasks, and sometimes using pre-VR training). Most of these researchers used their questions flexibly in response to respondent comments about their strategies, questions such as "How do you know that?" and "Why did you say that?" (Mareschal 2007; O'Malley et al. 1989; Vandergrift 1997, 1998b, 2003; Wagner 2008; Wu 1998; Yeldham and Gruba 2014, 2016). Such unstructured prompting, however, presents the possibility of missing some of the respondent's thoughts, as the line of questioning may lead the respondent away from thoughts heeded while they were listening to the text. To account for this – essentially, to 'mop-up' any remaining thoughts – Vandergrift (1997, 1998a, 2003) and Yeldham and Gruba (2014, 2016), added a final question such as "Anything else you'd like to say?" or "What else were you thinking?" when it was felt prudent.

In terms of the pre-VR instructions given to learners, many researchers have avoided highly-prescriptive directives, commonly just saying the passage would be paused at various junctures for the learners to say what they had been thinking while they were listening (Mareschal 2007; Murphy 1985; Vandergrift 1997, 1998a, 2003; Wagner 2008; Yeldham and Gruba 2014, 2016; Young 1997), and additionally, in Murphy (1985) and Yeldham and Gruba (2014, 2016) for the

---

1 While some of the studies cited here appear dated, they are still relevant to the current discussion of ways to conduct listener VRs.

learners to first say what they had heard. However, in many other studies, quite often those involving minimal researcher prompting, learners were provided with quite prescriptive directives, in particular to say how they had tried to understand the text, apparently to focus their report on the most pertinent aspects of the research (Bacon 1992; Goh 1998, 2002; Graham 1997; Graham et al. 2008, 2011; Gruba 1999; Laviosa 1991).

Only three studies, Buck (1990), Gruba (1999) and Yeldham and Chen (2014), appear to have examined VR mediation styles for L2 listeners. Buck (1990) compared the use of unstructured questions with structured (pre-determined) ones for Japanese EFL learners, concluding that the unstructured questions offered more flexibility to explore learners' processing, while the structured questions allowed better comparability between subjects' responses, but also pigeonholed their responses. Gruba (1999) then used Buck's structured questions in a pilot study with Australian learners of Japanese as a FL, but learner indifference to the rigid questioning led him to add some flexibility to the prompts later in his research. Finally, Yeldham and Chen (2014) compared unmediated and mediated VRs, finding the mediated VRs were favored by the learners, and tended to produce the best insights into their strategy use. The study, though, was greatly limited by the involvement of only three participants, each of whom were at different proficiency levels (ranging from lower-intermediate to upper-intermediate level); different texts were also used for each learner; and there were some errors in delivering the question prompts.

Consequently, because of the dearth of studies examining VR methods for researching L2 listeners, and the shortcomings of Yeldham and Chen (2014), the present study investigates various mediation styles and different pre-VR instruction directives. At its core, the study examines similar mediation styles to Yeldham and Chen (2014), specifically one mediated style using unstructured questions, another mediated style using similar unstructured questions as the first style but also concluded by a question asking learners if there was anything else they were thinking, and a third, unmediated style. The current study, however, employs a much larger number of learners than Yeldham and Chen (2014), and adopts a far more systematic approach in collecting and analyzing the data. In particular, the participants' listening proficiency level and the texts used in the study are much more tightly controlled.

Note that this comparison of the three mediation styles (styles 1–3) uses the same low-prescriptive pre-VR instructions for all three styles (asking learners to say what they heard, and *what they were thinking* while listening). It seems likely, though, that with an unmediated VR, in particular, one can more effectively elicit strategies data if the instructions beforehand are quite prescriptive. Therefore, a further component of the study adds such a high-prescriptive

pre-VR directive for an unmediated VR, asking learners to say what they heard, and *how they tried to understand*, to compare learner response to this directive with that of the less-prescriptive one.

4 Research questions

The research questions addressed in the study are:
(1) What are the learners' preferences between the three VR mediation styles, and why?
(2) What type of strategies data does each of the three mediation styles produce, and which style provides the most useful data for listening strategies research?[2]
(3) What type of pre-VR directive is more effective with an unmediated VR, a low-prescriptive one or a high-prescriptive one?

To address question 3, there were three sub-questions:
(a) Do the learners perceive any difference between the two types of directive?
(b) Do the learners feel the two types of directive differentially influence the VR in any way?
(c) Do the different directives elicit different strategies data from the learners?

In conducting the study it must be noted that it relied, rather subjectively, on the learners' perceptions of their behaviours to address Question 1 and also Questions 3 a and b. Such perceptions can lessen a study's reliability, although they are considered a necessary by-product of research conducted from a constructivist perspective (Lincoln and Guba 1985; Patton 2015). However, the research questions these perceptions aimed to explore were integral to the study, and the learners' views here complemented, and in the case of Question 3 also triangulated, the remaining questions in the study, examined from analyzing the participants' VRs. Consequently, various measures were taken to enhance the reliability of this perceptual data (which was gathered through semi-structured interviews, with these measures outlined shortly in the Method section). Note, too, that such individual learner perceptions had also been integral to the previous studies of VR mediation styles outlined above

---

2 Due to limited space in this article, another key issue from the research, relating to reactivity effects of the VRs, is examined in a separate paper.

(Buck 1990; Gruba 1999; Yeldham and Chen 2014), further endorsing their use in this current research.

5 Method

5.1 Participants

Seventeen 19-year-old Taiwanese university students from a freshman general English class, twelve females and five males, participated in the study. Ten were from a class taught by Author One, while the other seven were from a class taught by a colleague. The university stratifies its freshman English classes based on the learners' university entrance English exam results, and the two classes from which the participants were drawn were at lower-intermediate to intermediate level. Evidence of this proficiency level specifically for their listening was obtained through their scores on the listening sub-test of the Upper-Intermediate level version of the General English Proficiency test (GEPT). This is a standardized test used widely in Taiwan (Roever and Pan 2008), which was administered immediately prior to the research. The participants' results from this test are shown in Table 1, where their scores ranged from 20 to 30 out of a total of 45, with a mean score of 23.8, which translates approximately to an IELTS level of 5.[3]

Table 1: Participants and their GEPT listening test scores.

Learner	Listening test score[1]	Learner	Listening test score	Learner	Listening test score
1. Sal	20	8. Nina	22	15. Tom	19
2. Kev	21	9. Dave	22	16. Shirley	28
3. Dot	25	10. Di	25	17. Tina	28
4. Wen	21	11. Liz	23		
5. Cate	23	12. Bev	25		
6. Sam	21	13. Al	26		
7. Kim	30	14. Rob	26		

Note: [1]Max score of 45.

---

3 While not all participants were completely at the same level based on their GEPT scores, they did constitute a relatively homogeneous subset of the learners in the classes, where the range in test scores was 12–41.

Learners at this proficiency level were considered ideal for the study because they commonly use a wide range of strategies when engaging with texts (Vandergrift 1998b; Yeldham and Gruba 2016). This compares with high-proficiency listeners who tend to process texts more automatically, often with little need for strategies (Field 2008), and also compares with low-proficiency listeners, who commonly have great difficulty comprehending most texts and, according to Vandergrift (1998b), thus rely largely on guessing. Consequently, it was felt the current participants in the study would likely produce data from their VRs which was more suitable for the research than learners at other proficiency levels.

The main component of the study investigating the three mediation styles (Research Questions 1 and 2) involved the first 14 of the learners listed in Table 1. Then the second component comparing the pre-VR instruction types (Research Question 3) included three of these 14 learners, Bev, Al and Rob (with the reasons for their inclusion explained shortly) together with the three final learners listed in Table 1, Tom, Shirley and Tina.[4]

5.2 Data collection methods and procedures

The VR report styles used in the study are shown in Table 2: styles 1, 2 and 3 each used low-prescriptive pre-VR instructions, while unmediated style 4 was preceded by high-prescriptive instructions. The data for the study was gathered by conducting a VR followed by a semi-structured interview with each participant (Patton 2015). Collection of this data took place in Author One's office, and all the procedures were conducted in Chinese, by Author Two, to allow the learners to express their thoughts as accurately as possible (Cohen 2011). All procedures were audio-recorded for later translation and transcription. While the VRs were in progress, Author One operated the computer playing the listening texts, and other than welcoming the students to the room, generally remained unobtrusive to help reduce any possible learner anxiety caused by having two adults present. The VR session and interview lasted approximately

---

4 Originally 14 students were earmarked for the study. As the study progressed, though, it became apparent there was a need for extra learners from outside this original cohort to more effectively examine Research Question 3: this was ostensibly to add participants to the study who were uninfluenced by earlier exposure to mediation styles 1 and 2 (see later in the paper why this was a concern). Unfortunately, though, limitations in time and participant availability meant that only a small number of these extra learners were added.

Table 2: Verbal report styles used.

Style	Characteristics of style	Main rationale for style
Main mediation styles used		
Style 1: Unstructured mediation	Questions asked flexibly in response to listener comments (e. g., "Why did you say __?"; "Why did you think __?"; "How did you know __?").	Allows researcher flexibility to examine emerging areas of interest. Prompts used mainly at times of learner comprehension difficulty, when strategic processes most likely to be consciously accessible for report (Færch and Kasper 1987).
Style 2: Partly-structured mediation	Abbreviated version of style 1, but every interviewer-respondent exchange concluded with, 'Was there anything else you were thinking that you forgot to tell me?'	Final question added to elicit any further thoughts the respondent may have neglected to report earlier in the exchange (Yeldham and Gruba 2016).
Style 3: Unmediated	No questions asked. Instead, VR focused by low-prescriptive pre-sessional instructions (used for Styles 1–3) asking the learner to say 1) what they heard, and 2) *what they were thinking* while they were listening.	Questions not used, to avoid interfering with respondent's cognitive processing (Ericsson and Simon 1993). Low-prescriptive pre-sessional instructions used to avoid influencing learner response.
Additional pre-VR instruction style for unmediated VR		
Style 4 Prescriptive pre-sessional instructions	No questions asked, but VR focused by prescriptive pre-sessional instructions asking the learner to say 1) what they heard, and 2) *how they tried to understand* while they were listening.	Prescriptive pre-sessional instructions used to focus the report on area of research interest (Goh 1998).

one hour with each participant, and each session was punctuated by occasional short breaks to allow the learners to refresh themselves mentally. Appendix A outlines the task directions and sample questions used while collecting the data from the VRs; Appendix B outlines sample questions used while collecting the interview data.

5.3 Verbal report procedures

Immediately before their VR, the participants practiced verbalizing their thoughts with two non-listening tasks (Ericsson and Simon 1993). The first was

a mathematics task requiring the learners to multiply 32 by 12, and in the second, they had to describe what they 'saw' as they imagined entering and walking through their house. Listening tasks were not used at this practice stage, mainly to avoid the possibility of alerting the learners to any listening strategies we considered desirable. As Goh (1998) has pointed out, such awareness may lead learners to later use, or perhaps report using, these strategies to please the researcher. After the warm-up task, the participants were then clearly instructed (Cohen 2011) what to do during the VR, which was read to them from the script shown below:[5]

> Now we're going to play some listening passages to you. We will regularly stop each passage after every sentence or two, and each time we stop it, we'd like you to tell us *what you heard*, and also *what you were thinking while you were listening*. Please tell us everything that was in your mind; please give us as full and as accurate an account of your thoughts as you can.

Each text was played without first telling the listener its topic, to force the learners to hypothesize the text's content based on what they heard (Vandergrift 2003); to assist this hypothesis development, texts were used where the topic was indicated early in the passage. We stopped the texts for the listeners to report their thoughts after one long sentence or two shorter ones. We considered segments of this length short enough to avoid overly taxing the listeners' WM (exceeding WM limits could have pushed them to infer their strategic processes or to access past experiences in their VR), and also long enough to prevent them from recalling the segment of text verbatim via echoic memory.

Guided by the verbal report literature, certain restrictions were placed on the use of prompts through the study. First, questions were limited to asking respondents to describe their behaviors; they were not asked to *explain* why they used certain behaviors ("Why did you guess this?") to reduce the possibility of inducing changes in processing from the greater level of rationalization involved (Ericsson and Simon 1993). The interviewer also avoided: (a) asking leading questions that could have influenced the subjects' responses (Graham 1997); (b) asking questions that may have highlighted strategies for the subjects to use later in the task, such as "What do you predict the speaker might say next?"; (c) asking the learners what they thought the text was about while it was

---

5 We also specified any change in directions before each modification in style: e. g., "From now on we will ask you some questions [...]", "This time [...] we want you to say how you tried to understand."

in progress, as this may have motivated them to monitor their comprehension more than usual; and (d) mentioning any preceding text content in the questions asked, as this could have helped the learners better comprehend subsequent content (Rost 2001).

5.4 Interviews

The semi-structured interviews were guided by an interview schedule (Patton 2015; see Appendix B). For the first component of the study, investigating the 14 learners' views of the three different mediation styles, the learners were asked: how they felt about each style, including which one they liked the most and why, and which one they liked the least and why. For the second component of the study, comparing the two pre-VR instruction styles, learners were asked if they perceived a difference between the two styles, and if they felt any different effects on the VR from the styles. To enhance the reliability of the data all the interviews were conducted immediately after the participants' VR, to maximize the learners' ability to recall their thoughts about the relevant issues. Also, prior to asking the learners their thoughts about the various styles, the interviewer clearly outlined to them the characteristics of each style and then checked to make sure they were clear about which style was which.

5.5 Texts used and text/learner configurations

There were five listening texts used throughout the study, and three more were added for the study's second component. All the texts had been used in Author One's previous research (Yeldham 2009). They had been trialed for that research, and found suitable for Taiwanese learners at the same proficiency level as the current participants, both in terms of their language difficulty and their topic familiarity.

The five listening texts used through the study were sorted into three text blocks each consisting of seven text segments (with each segment being one long sentence or two short sentences long). Text block A consisted of a single, rather lengthy text, *Sunbathing in the West and Taiwan* (seven text segments). Text block B included two shorter texts, *Text Messages* (four segments) and *The Dog and the Duck* (three segments). Text block C also included two short texts, *The Hotel* (three segments) and *Adventure Activities* (four segments).

As shown in Table 3, to balance out possible presentation order effects, for seven of the learners, mediation style 1 was used first, followed by style 2 then style 3, then this order was reversed for the other seven learners. To balance possible text effects, the order of the text blocks A, B and C was rotated between the learners: one learner would receive the texts in the order A, B, C; the next learner in the order B, C, A, and the next, C, A, B, with this pattern recycled for all participants.

Table 3: Mediation style comparison: Mediation style and text block, by learner.

Mediation style	Order of text blocks						
	Sal	Kev	Dot	Wen	Cate	Sam	Kim
1	A	B	C	A	B	C	A
2	B	C	A	B	C	A	B
3	C	A	B	C	A	B	C
	Nina	Dave	Di	Bev	Al	Liz	Rob
3	C	A	B	C	A	B	C
2	A	B	C	A	B	C	A
1	B	C	A	B	C	A	B

Of the three additional learners recruited specifically to compare the pre-VR instruction styles (Tom, Shirley and Tina), for each of these two styles (styles 3 and 4) each learner also listened to one text block. As shown in Table 4, these passages were recycled between the learners, in similar fashion to the comparison of styles 1, 2, and 3 (in Table 3). Three learners from the mediation style comparison, Rob, Al and Bev, were also included in this pre-VR instruction style comparison. All three were deemed suitable because the unmediated style 3 was the first style they had been

Table 4: Pre-VR instruction comparison: Mediation style and text block for three of the learners*.

Mediation style	Order of text blocks		
	Tom	Shirley	Tina
3	A	B	C
4	B	C	A

Note: *The other three participants in this component of the study listened to other texts for style 4.

exposed to in the mediation style comparison (had they first been exposed to the mediated styles it would have alerted them to the main aim of the study, that of eliciting their strategy use, before they were then exposed to style 3). Consequently, they were judged as being relatively capable of comparing styles 3 and 4 (though perhaps not perfectly, as they had also experienced the mediated styles after style 3). In using style 4 for these three learners, all listened to 6 segments from *Welfare in Taiwan and the West*. Bev had to leave after a short time, so she listened to this text only. Al had considerably more time available so he also listened to *The relationship between teachers and students in Taiwan* (6 segments). Rob had less time available than Al, so he listened to *Courtyard houses* (4 segments). This variation lessened comparability across the learners, but we had to adjust on the run as the time constraints presented themselves.

5.6 Data analysis

To analyze the interview data for Research Question 1, we examined the learners' interview answers regarding their mediation style preferences. The answers were quite explicit so the analysis of this low-inference data was fairly straightforward.

In analyzing data for Research Question 2, we heeded advice from Cohen (2011) and Kasper (1998) to enhance the validity of the findings by reporting and displaying both comprehensive data (within practical limits, though, given the massive dataset) and also more fine-grained results selected from the data. For half of the learners (those being the first seven shown in Table 3, for whom style 1 was the first style used) we coded all of the 49 segments (7 from each learner) elicited from each of the three styles. We also coded the 49 segments elicited from style 3 for the remaining seven learners, this style being the first style they were exposed to. We coded these further segments from style 3 because we suspected a different outcome from this style 3 depending on when it was used relative to the mediated styles. When used after exposure to the prompts of the two mediated styles the learners were more likely to be aware of what behaviors they were expected to report than when style 3 was the first style used.

We used this data to compare the number and types of strategies elicited from the two mediated styles with that of the unmediated style 3. In particular, we used paired samples t-tests to compare the number of strategies elicited from the 7 learners for whom strategies were coded from all three styles, comparing between these learners the tallies for style 1 with style 3, and for style 2 with style 3. While the learner sample size here was small, the t-tests were permissible because the data was normally distributed and demonstrated homogeneity of variance.

Additionally, to compare and examine strategy use elicited from all three styles in more detail, for two of the participants, Di and Bev (for whom style 3 was the first style used, and thus untainted by prior exposure to mediated styles), we coded their data elicited by all three styles (14 segments from each style, in total). We chose these two learners because we felt their response was representative of the differences in strategies elicited from the three styles. Also, to further illustrate the differences in data elicited by the mediated style 1 and the unmediated style 3 we also briefly qualitatively analyzed transcribed VR excerpts from one of these two learners, Bev. Finally, to compare strategies elicited between the two mediated styles 1 and 2 in more quantitative terms, we compared 63 segments from each of these 2 styles (the 49 segments from the 7 learners mentioned in the paragraph above, and 14 segments elicited from Di and Bev). We compared these strategy totals across the nine learners using paired samples t-tests (permissible because this data was found to be normally distributed and showed homogeneity of variance). Because we used multiple t-tests through the study, to reduce the possibility of type-1 error, a Bonferroni adjustment set the significance level at $p < 0.013$.

Additionally, to further examine the impact of the final, structured question of style 2, we categorized all 96 responses (of the potential 98 responses, the interviewer forgot to ask it twice) elicited by that question. Figure 1 summarizes the data analyzed to address Research Question 2.

Style 1	*Style 2*	*Style 3*
49 segments coded (all 7 segments from the 7 learners for whom this was the first style used)	49 segments coded (all 7 segments from the same 7 learners as for style 1)	49 segments coded (all 7 segments from the same 7 learners as for styles 1 and 2)
All segments from all 3 styles coded and presented in detail from two of the learners[1]		
	Final question from style 2 was categorized for all segments from all 14 learners.	49 segments coded (all 7 segments from the 7 learners for whom this was the first style used)

Figure 1: VR data coded to address Research Question 2.
[1]Also VR data from one of these two learners was presented with a brief qualitative analysis of that data.

For Research Question 3, we examined the interview answers by all six of the learners involved. We also coded the strategies elicited using styles 3 and 4 for the three learners who were not involved in the mediation style comparison (the other three were excluded from this analysis because their style 4 VRs would have been greatly influenced by their prior exposure to mediated styles 1 and 2).

Author One coded the VR data, guided by strategy classifications used in Yeldham and Gruba (2016), with the addition of some other categories of strategies and behaviors observed in the data. To establish inter-rater reliability, Author Two then independently coded sections of the data first coded by Author One, including: (a) 21 segments elicited through style 1 (just over 40 % of the 49 segments); (b) 21 segments elicited through style 2 (just over 40 % of the 49 segments); (c) 21 segments elicited through style 3 (just over 20 % of the total 98 segments); (d) and 21 segments (just over 20 %) of the dataset elicited by the final question of style 2; (e) all 28 data segments elicited by styles 1 and 2 from the two selected learners; and (f) one third of the data elicited from the comparison of styles 3 and 4. Inter-rater reliability, established by calculating the number of agreements, divided by the total number of agreements plus disagreements (Miles and Huberman 1994), was respectively, (a) 0.81, (b) 0.84, (c) 0.86, (d) 0.9, (e) 0.82, and (f) 0.86.[6]

6 Results

6.1 Research Question 1

Research Question 1 was: *What are the learners' preferences between the three VR mediation styles, and why?*

From the interview data, Table 5 shows each learner's preferred mediation style, indicated with a check, and their least favored style, indicated with a cross. Style 1, unstructured mediation, was the most favored style, with 8 of

Table 5: Learner mediation style preferences.

Style	Sal	Kev	Dot	Wen	Cate	Sam	Kim	Nina	Dave	Di	Bev	Al	Liz[1]	Rob	Total
1			✓	✓	✓		✓	✓	✓	✓	✓				+8
2	✓	✓				✓						✓		✓	+5
3	×	×	×	×	×	×	×	×		×	×	×	×	×	−14

Note: Most-preferred style ✓; Least-preferred style ×.
[1]Liz had no preferred style, but suggested alternating between styles 1 and 2 might be best.

---

[6] A figure of.8 or above is commonly considered a respectable reliability figure for VR coding (Green, 1998).

218 — Michael Yeldham and Rainbow Tsai-Hung Chen

the learners preferring it. This was followed by Style 2, which was favored by a further 5 learners. One learner, Liz, had no preferred style but suggested conducting VRs by alternating between styles 1 and 2. The least preferred style among the learners was unanimously the unmediated style 3.

As summarized in Table 6, the most frequent reasons why learners liked style 1 were that it helped to guide them what to report on (5 comments), elicit more information (3 comments) and to remind them what they had been think-ing (2 comments). Note that these comments endorse many of the reasons for using researcher prompts outlined earlier in the article by Cohen (2011), Jourdenais (2001) and others. These learners also expressed a less favorable opinion of style 2, with six of the seven learners saying this was because they found the earlier unstructured questions elicited all their relevant thoughts, rendering the final structured question unnecessary. Two of the learners also felt that the final question made them feel pressured from having to constantly try to answer it.

Table 6: Learner reasons for preferring Style 1 (and not Style 2).

Name	Why prefer Style 1	Why not Style 2
Bev	Qs help guide me to say what thinking processes to report on	Earlier Qs elicit all information, so final Q unnecessary
Dave	Qs help to guide response	–
Nina	Qs guide me to remember content; Won't have to keep stretching to explain my thinking method	Earlier Qs elicit all information, so final Q unnecessary
Dot	Qs remind me what I was thinking.	As above
Cate	Qs draw attention to important parts, thus response more detailed.	As above (+ felt pressure)
Di	Qs elicit more response	As above (+ felt pressure)
Wen	Will think of precise answer to Q.	As above
Kim	Qs are based on what I've already said. Like an extension that helps my thinking.	As above

Among the five learners who preferred style 2 (see Table 7), three claimed the final question helped remind them of any thoughts they had missed reporting, which was consistent with the reason proposed for adding such a question (Vandergrift 2003; Yeldham and Gruba 2016). The other two learners (Sal and Kev), however, suggested the style may have given them some assistance: for Sal

Table 7: Learner reasons for preferring Style 2 (and not Style 1).

Name	Why prefer Style 2	Why not Style 1
Rob	Qs help guide what thinking processes I should report on; final Q helps me make sure I didn't miss anything I want to say.	–
Al	Qs give guidance; final Q helps me make sure I didn't miss out anything.	–
Sam	Qs give guidance; final Q helps me realize if I forgot or mistook something	Less complete than style 2
Sal	Final Q makes me reflect on everything I've heard	As above
Kev	Final Q gives another chance to think	–

this was in making her "reflect on everything I've heard", and for Kev, it gave him "another chance to think".[7]

All 12 of the learners considered style 3 their least preferred style. Table 8 outlines the reason, or reasons, for this by each learner. Those for whom it was the first of the three styles used are shown separately from those for whom it was the last style used. This is because the latter group would have had a clearer idea of what listening processes to report than the former group after having just previously received the question prompts of styles 1 and 2.

The results showed that when it was the first style used, 5 of the 7 learners in this group complained that they didn't really know what to say. One more even felt such pressure at having to do everything by himself that he compared the experience to being interrogated! When the style was the last of the three used, the main complaints, by 6 of these 7 learners, were that it was difficult for them to remember their thoughts or that they felt there were many thoughts left unreported. So even having a good idea of what was required for the task, it seemed for this group that the act of reporting what they understood and how they tried to understand it was too much for them to do on their own.

6.2 Research Question 2

Research Question 2 was: *What type of strategies data does each of the three mediation styles produce, and which style provides the most useful data for listening strategies research?*

---

7 These two final explanations, along perhaps with the final comment in Table 6, imply some degree of positive reactivity. This is taken up in a later article focusing on reactivity from this research.

Table 8: Learner reasons for Style 3 being the least preferred.

Learner	Reasons why
When it is the first style used	
Nina	Couldn't think of much to say, and there's no one to guide me
Dave	Have to do everything by myself; Feels like I'm being interrogated
Di	Don't know what to talk about
Bev	Don't know what to talk about
Al	Don't know what to talk about
Liz	Too much like a test
Rob	Feels like I'm talking to myself; I don't know the purpose of saying those things.
When it is the final style used	
Sal	Easier to forget things
Kev	Nothing was asked, so I'll talk less
Dot	Tend to forget information
Wen	I don't really think: my head will be empty
Cate	Many things will be lost: I won't think of so many things to talk about
Sam	Have to do everything by myself
Kim	I only say what I think is useful, so details might be left behind

One clear result after examining the VR data, was that the data elicited from the unmediated style 3 generally differed from that elicited by the two mediated styles. This difference was particularly clear for data elicited from style 3 when it was the first style used.

Table 9 summarizes the data elicited from the first 7 learners from all three styles (49 researcher-respondent exchanges from each style) when style 1 was the first style used with these learners, and also in the right column, an additional 49 such exchanges from the other seven learners for style 3 when *it* was the first style used. The main difference in the results is that far more strategies (both lower-level and higher-level ones) were elicited from mediated styles 1 and 2 than from unmediated style 3. The total number of strategies elicited from style 1 was 72, from style 2 it was 68, and from style 3, when used with these same learners, a far fewer 42. The paired samples t-tests showed significant differences here at the alpha level of .013 across these 7 learners: differences between style 1 ($M = 10.29$, $SD = 2.14$) and style 3 ($M = 6.0$, $SD = 1.53$) ($t = 5.12$, $p = 0.002$), and also between style 2 ($M = 9.71$, $SD = 2.14$) and style 3 ($t = 4.44$, $p = 0.004$). When style 3 was the first style used (with the other seven learners, who were thus uninfluenced before-hand by mediated VRs), even fewer strategies were elicited, at 32. Clearly this shows that the use of questions probing learner processing (styles 1 and 2) garnered more information about how learners engaged with the text than when such questions

Table 9: Summary of listening strategies (and other behaviours) elicited from each style.

Strategy/behavior used	The first 7 learners			The second 7 learners
	Style 1 (1st style used)*	Style 2 (2nd style used)	Style 3 (3nd style used)	Style 3 (1st style used)
Say words interpretation based on[1]	39	38	7	15
Link key words[2]	3	–	10	5
Translate[3]	1	1 (3)**	2	3
Lower-level strategies total	**43**	**39 (41)**	**19**	**23**
Predict[4]	2	1 (2)	2	–
Guess/infer[5]	6	6	5	2
Elaborate[6]	14	6 (7)	3	5
Questioning elaborate[7]	1	3	3	–
Use imagery[8]	1	1	1	1
Attend selectively[9]	–	2 (4)	–	1
Monitor comprehension[10]	5	4	9	–
Higher-level strategies total	**29**	**23 (27)**	**23**	**9**
Total strategies elicited	**72**	**62(68)**	**42**	**32**
Comment on content	1	1	–	4
Interpret text only	2	8	11	11
Fail to interpret text (too hard)	1	1	–	2

Note: * Each of these four columns of summarized data in the table came from 49 text segments; ** Figures in parentheses indicate addition of strategies elicited from the final question of style 2 (Was there anything else …?)
[1]Indicate words used to interpret text; [2]Mention that he/she pieced words together to interpret text; [3]Change information from L2 into L1; [4]Predict words/information; [5]Hypothesize meaning of words/information; [6]Use existing knowledge to fill in missing information; [7]Use questions and knowledge of the world to brainstorm possibilities; [8]Use mental image to assist comprehension; [9]Focus on or ignore parts of text; [10]Check, verify or correct one's discourse level interpretation.

were absent (style 3). In turn, after previous exposure to the mediated styles, when style 3 was used as the last style presented to the learners it elicited more strategies than when it was the first style used. This was especially obvious in terms of comprehension monitoring (style 3 used first, 0 instances, style 3 used last, 9 instances). Also apparent from Table 9 is that in a number of the responses elicited from style 3, the learners did not mention their strategy use at all (11 times each when the style was used first and last), simply interpreting the section of text they had heard.

Not reflected in Table 9 is the further difference between the unmediated and mediated styles in terms of reporting on how the strategies were used. By comparison with styles 1 and 2, style 3, especially when it was the style used first, often elicited vague, or generalized, descriptions of strategies use, with learners simply saying "I linked the words", "I translated", or "I guessed the meaning", with no evidence to show that the strategy was in fact used, or little indication of what lower-level or higher-level information was used in the strategy or the specific comprehension problem the strategy was used to address.

The following interviewer-learner exchanges with Bev (B) for two texts illustrate the differences between style 3 (which was the first style used with Bev) and style 1; Bev's response was chosen because it typifies this difference (note that less-relevant data has been removed from the examples due to space concerns). Through the first text, *The Hotel*, using style 3, after giving her interpretation of each text segment, Bev simply outlines the strategies she says she used, with no insight into when they were used in the listening process, how she applied them or which information she drew from (e. g., "I'd try to link them [the words] together"; and "I just took what I heard and then translated it into Chinese").

Bev: Style 3 (Unmediated)

1-1 [Text]

B: [B's Interpretation] Thinking process ... well ... I took what I heard and I tried to see if there were any words that I've learnt before ... that's all. And then if there were any, I'd try to link them together. Then if I hadn't heard them before ... I'd remember the key words.

1-2 [Text]

B: [B's Interpretation] Thinking process ... it's also the same ... if I heard some familiar words, I would put them together.

1-3 [Text]

B: [B's Interpretation] Then ... thinking process ... I just took what I heard and then translated it into Chinese ... didn't really have any thinking process.

By comparison, through the second text, *The Dog and the Duck*, style 1 yields much more information. The first exchange shows how interviewer (I) questioning elicits from Bev how she elaborates using her rhetorical knowledge to help strengthen her interpretation of the text. In the second exchange, the interviewer's question then

elicits how Bev knows her change in mental model, gleaned from monitoring her comprehension, is a correct one. Finally, the question in the third exchange draws out the specific bottom-up information Bev used to help her interpret the text.

Bev: Style 1 (Mediated)

2-1 [Text]

B: [B's Interpretation] [...] I think it's a news report. [...]
I: Why do you think that it's the news?
B: Because the way the sentence started sounds like some news. In the news, they will first tell you where something happened. And then it will summarize in a short segment what the news is about. And then after that, it will describe more about why this happened. Isn't the news all like this?

2-2 [Text]

B: [B's Interpretation]. Yeah. I got it wrong, the last sentence. [...] So ... because I understand this sentence, that's why I know that I got the last sentence wrong.
I: How do you know this sentence is the right one instead of the first one?
B: Because more was said in this sentence, but in the previous sentence, only one statement was made. Yeah, and so I'll be a little ... if I missed the sentence, then I miss the whole thing. In this section, where there are more sentences, and I was able to make the right judgements.

2-3 [Text]

B: [B's Interpretation]
I: How did you know that it said 'if the dog leaves, the duck will be sad?'
B: When the dog leaves ... there's the word 'sad', and then also the word 'quack' meaning it keeps making noises. Only until the dog returns will the duck stop make noises and be sad because of the word 'until'.

Comparing the strategies elicited from two of the learners (Di being one, and Bev the other) across all three styles (Table 10), confirms the various findings outlined above. From both learners' responses combined, the unmediated style 3 only elicits one strategy of selective attention and one guess, with the rest, as shown in Table 10, generally being unsupported claims of linking key words and translation (e. g., "I linked key words"). By comparison, style 2 elicits a wider range of strategies (including 2 guesses, selective attention, imagery and elaboration), as does style 1 (6 guesses, 2 elaborations, selective attention and comprehension monitoring). For these two mediated styles, additional evidence

Table 10: Strategies and other behaviors elicited from two learners (Di and Bev).

Seg.	Di (style 3)	Di (style 2)	Di (style 1)	Bev (style 3)	Bev (style 2)	Bev (style 1)
1	Says words from text (does not Int text)[1]	Links key words to Int text	Says words Int based on	Says she linked key words[2]	Says words Int based on	Says words Int based on
2	Says she linked key words	Uses imagery	Identifies problem word, tries to guess it	Says she linked key words	Guesses meaning, and explains how	Links key words
3	Says she linked key words	Says she linked key words	Says she linked key words	Says translated	Says words Int based on	Says she guessed the meaning
4	Says she linked key words; says translated	Says words Int based on	Elaborates; Guesses meaning, and explains how	–	Says words Int based on	Guesses, and says words guess based on
5	[Comments on content]	Says words Int based on; Elaborates[3]	Guesses word meaning, and explains how	–	Says she linked key words	Elaborates
6	Attends selectively	Says words Int based on	Says words Int based on	–	Attends selectively	Monitors, and changes mental model
7	Says words Int based on	Guesses word meaning, and explains how	Guesses word meaning; Attends selectively	Guesses word meaning, and explains how	Says words Int based on	Says words Int based on

Note: [1]The respondents interpreted (Int) all other text segments shown in this table; [2]Bev's extended interviewer-learner exchanges shown earlier come from exchanges 1–3 from style 3 and 5–7 from style 1; [3]This elaboration was elicited from the final question of style 2 (and is the only strategy shown in this table which was).

(what information they used from the text, and so forth) was also commonly provided by the learners to confirm the strategies they reported.

Finally, comparison of the two mediated styles, styles 1 and 2, in more quantitative terms was determined through strategy totals for each style from the 49 coded segments shown in Table 9 added to the strategy totals for these two styles elicited from Di and Bev (from Table 10). These yielded 88 strategies elicited from style 1, and by comparison, 76 from style 2 when not including strategies gathered from the final VR question (this difference was expected because fewer unstructured prompts were used here with style 2 than with style 1), and 83 strategies from style 2 when these additional strategies were added – still short of the total number garnered from style 1. The paired samples t-tests showed no significant differences across the individual learners in comparing style 1 (M = 9.78, SD = 2.17) with style 2 when strategies from the final VR question were not included (M = 8.44, SD = 2.0) (t = 1.52, p = 0.169), nor comparing style 1 with style 2 when strategies from the final VR question were included (M = 9.22, SD = 2.11) (t = 0.73, p = .489).

Further, underlining the lack of advantage gained from the final question of style 2, of the 96 times the question was asked (for all learners), the answer was simply "No" on 61 occasions (approximately two-thirds of the responses). A further 11 times learners just reiterated what they had said earlier, occasionally in slightly more detail, and 9 times learners made miscellaneous comments, chiefly commenting on the text content or mentioning comprehension difficulties they were having. Only twice did learners add to their interpretation of the text. The remaining 13 responses elicited strategy use, but often of a generalized nature, much like the responses to style 3 (e. g., "I translated"). The only really useful contribution came from Nina describing in consecutive exchanges how she was recycling the meaning of parts of the text in her mind to try to understand it.

6.3 Research Question 3

Research Question 3 was: *What type of pre-VR directive is more effective with an unmediated VR, a low-prescriptive one or a high-prescriptive one?*

To address this question, the three sub-questions were:
(a) Do the learners perceive any difference between the two types of directive?
(b) Do the learners feel the two types of directive differentially influence the VR in any way?
(c) Do the different directives elicit different strategies data from the learners?

Firstly, as shown in Table 11, four of the learners did not notice any difference between the two styles. Tom and Shirley, explained that they were alerted to the need to report how they tried to understand the text in style 3, from having to solve the prior warm-up tasks. Another of the four learners, Al, said he just knew as well, but without realizing why. Consequently, all four learners felt they responded equally with both styles, and felt they reported much the same data from each.

Table 11: Learners' comparisons of styles 3 and 4.

Learner	Perceive any difference between styles?	Feel styles differentially influence the VR?
For 3 learners from mediation style comparison		
Bev	No. (No reason given.)	Felt her answers were "more or less the same."
Al	No. Said he just knew that we wanted him to report how he tried to understand in style 3.	Says his reports were the same for both styles.
Rob	Yes. Felt Style 3 better.	Felt pressured to restrict what he reported with the more specific directive of style 4.
For additional 3 learners, not from mediation comparison		
Tom	No. Realized from warm-up problem solving task that requirement of style 3 was to report how he tried to understand.	Because he knew what was required of style 3, the style did not impose greater stress on him than style 4.
Shirley	No. Realized from mathematics warm-up problem solving task that requirement of style 3 was to report how she tried to understand.	Felt that thinking how to answer the directive of both styles interfered with her listening equally.
Tina	Yes. Felt style 4 better. Said warm-up task did not alert her to "real" requirement of style 3.	Felt more relaxed with style 4 as it offered her better guidance on what to report.

Another learner, Rob, felt the more-prescriptive style 4 restricted what he wanted to say more than style 3, explaining style 4 might encourage him "to deny away my thoughts," also making him feel more pressure than with style 3. One observation here about Rob, though, is that throughout his VR for styles 1–3 he reported less than most of the other learners on how he overcame listening problems and more on describing images he had while listening and commenting on the content. Given this context, it seems understandable he disliked the more prescriptive style 4.

The final learner, Tina, preferred style 4, saying its directives were "more specific, concrete, plain, and straightforward." Consequently, she said this style helped her to relax more than style 3.

In examining the VR data for the three learners included in the research solely for this component of the study, Tom, Shirley and Tina, firstly for Tom, styles 3 and 4 produced virtually the same information. For each segment, after first giving his interpretation of it, he then commonly just verbalized the words he understood to help him reach this interpretation (see Table 12). Shirley's response was also quite similar for both styles: beyond verbalizing the words she used to interpret the text, and saying she translated, style 3 elicited one instance of selective attention and one guess, while style 4 elicited from her two guesses and one prediction. For Tina, style 4 (her favored style) produced marginally more information than style 3. Both styles elicited from her that

Table 12: Strategies elicited in comparing unmediated styles 3 and 4.

Seg.	Tom		Shirley		Tina	
	Style 3	Style 4	Style 3	Style 4	Style 3	Style 4
1	–	Says words Int based on[1]	Says words Int based on	Says words Int based on	Says she linked key words	Says words Int based on
2	Says words Int based on	Says words Int based on	Says words Int based on	Says words Int based on; Says translated	Says she linked key words	Says translated
3	Says words Int based on	Says words Int based on	Says words Int based on	Says words Int based on	Says translated	Elaborates; Says translated
4	–	Says words Int based on	Says translated	Says translated	Says translated	Says translated
5	Says words Int based on; Says translated	Says words Int based on	Says words Int based on	Predicts; Says translated	Says translated	Says translated
6	Says words Int based on	Says words Int based on	Says words Int based on; Selective attention	Guesses meaning, and explains how	Says translated	Says she linked key words
7	Says words Int based on	Says words Int based on	Guesses meaning, and explains how	Guesses word meaning, and explains how	Says translated	Elaborates

Note: [1]The respondents also interpreted (Int) all text segments shown in this table.

she linked key words or translated, but with Style 4 she also explained twice that she elaborated using prior knowledge.

So in summary, the comparison of styles 3 and 4 by this small learner cohort generally showed little difference in their perceived responses to the two styles. Both styles also produced similar strategies data, with the exception of one learner for whom the more-prescriptive style 4 tended to elicit a slightly wider number of strategies than style 3.

7 Discussion and conclusion

The results from the study strongly endorse the use of researcher questions to probe learners' thoughts in VRs. The findings thus support a constructivist position (Afflerbach and Cho 2009; Cohen 1998, 2011; Vandergrift 1998a) over that of a less-interventionist information processing one (Ericsson and Simon 1993; Fox et al. 2011; Green 1998) in indicating that for a multidimensional task such as listening comprehension, it appears best to employ questions to help elicit the respondent's VR data.

In advocating mediated VRs, an unstructured style with questions used flexibly in response to learner comments (Vandergrift 2003), that of style 1, seems the best approach. In the study it was favored by the majority of the learners, and together with the partly-structured mediation of style 2 produced more useful data for strategies researchers than the unmediated VRs of style 3, both by eliciting more strategies, and greater insights into how, why and when the strategies were used. The partly-structured mediation style was also considered the most useful style by some learners, although its final question added to elicit any final respondent thoughts (Vandergrift 2003; Yeldham and Gruba 2016) was generally found to be unnecessary, as it often produced little useful information. Some learners also felt pressure from having to constantly answer this final question.

The study also emphasized how unsuited an unmediated style is for listener VRs: unmediated style 3 was considered the least popular style by all participants. For learners who encountered it as the first VR style used, they disliked its lack of guidance (Cohen 2011). Indeed, these learners' verbal reports elicited few insights into their strategies use. However, for those learners who encountered this style after first encountering the two mediated styles, and thus had a better idea of what to report, their reports did provide some useful accounts of their strategies. This perhaps suggested the practical possibility of using unmediated VRs after prior learner training with mediated ones (O'Malley et al 1989), an approach which may appeal to researchers concerned about prompts intruding

on the learners' report (Green 1998). However, the data elicited when this unmediated style 3 was used last still did not hold up to that of the mediated styles. In addition, even when used last, the learners disliked this style's lack of guidance, with most feeling that its absence of questions made it hard for them to remember their thoughts or left many of their thoughts unreported (Cohen 2011; Jourdenais 2001). Finally, in the study it was anticipated that directions outlining more specifically what was required in an unmediated VR – that is, asking the learners to explain how they *tried to understand the text* (Goh 2002; Graham et al. 2011) – may have improved the effectiveness of such a VR. However, while this more-prescriptive style 4 did elicit slightly more strategies than style 3 from one learner, overall the difference between the two styles was minimal. In addition, most of the learners who experienced both styles felt that the different forms of directive from each produced a similar response from them in their VR, with some realizing from the warm-up activity that being asked to express their thoughts while listening really meant to explain how they tried to understand the text. Consequently, there seems little advantage in employing such a high-prescriptive directive.

Besides these suggestions stemming from the research, one further practical implication from it is that the perceived need for responsive researcher questioning underscores the necessity for conducting VRs one learner at a time. Had unmediated verbal reporting been found a more useful approach, it could have signaled the appealing possibility of conducting VRs for listeners en masse, conducted by recording large groups of students' responses concurrently in a computer room or a language lab (Cohen 2011), thus saving substantial data collection time. Alas, though, this does not seem feasible in light of the results.

Naturally there were limitations to the study. One concerns, to some degree, the participant sample size, and the difficulty of generalizing findings from these learners to a wider population. Consumers of qualitative research, however, acknowledge the practical constraints imposed in gathering often large amounts of data from the participants, and thus see value in the depth to which the participants are examined. For transferability to other settings, other factors are considered important, too, including detailed descriptions of the research procedures and the tying of components of the research to existing theories (Guba and Lincoln 1985), factors that we observed keenly in our study. In fact, we would argue that the number of participants we used to examine Research Questions 1 and 2 (14 learners), was quite large for a qualitative study, and that this enhances the transferability of this aspect of our research. With only 6 participants used to examine Research Question 3, though, we are less confident here. Another limitation to the study was that the participants were all Taiwanese EFL learners. Nevertheless, we think that such

learners share many characteristics with other Confucian-heritage Asian learners (Kim 2002), and thus feel that the findings may be extrapolated to intermediate-level learners within that wider context. One proviso, though, is that one must be mindful that the results are based on the use of a specific set of monologic texts. The relatively large number of these texts used, and their variety of topic and so forth, obviously increases the relevance of the study elsewhere, although there are no claims that the findings apply to other forms of listening such as interactive listening.

In conclusion, in suggesting a mediated method for L2 listening researchers, while the unstructured style 1 seemed the most effective, we feel reluctant to entirely dismiss the use of a concluding question to probe for any remaining respondent thoughts. In our study, for style 2 such a question was added to conclude *every* interviewer-respondent exchange. However, it could be that the question is a very useful one when used judiciously at certain times through the VR. Indeed, one learner did favor alternating between the unstructured style 1 and partly-structured style 2. Perhaps future research could investigate when such a final question is of most use during a VR. Finally, as the study examined Asian learners only, future research could also examine learners from non-Asian backgrounds. This appears useful in light of Kim's (2002) research showing that Asian-Americans are more reluctant to express their thoughts verbally than European-Americans, as it appears likely that learners from other cultural backgrounds may respond differently to VR styles such as those presented here.

References

Afflerbach, P. & B. Cho. 2009. Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. Israel & G. Duffy (eds.), *Handbook of reading comprehension research*, 69–90. Mahwah, NJ: Erlbaum Associates.

Bacon, S. M. 1992. Authentic listening in Spanish: How learners adjust their strategies to the difficulty of the input. *Hispania* 75. 398–412. doi.org/10.2307/344077

Bowles, M. A. 2010. *The think-aloud controversy in second language research*. New York, NY: Routledge.

Buck, G. 1990. *The testing of second language listening comprehension*. Unpublished Doctoral Dissertation. University of Lancaster, UK.

Cohen, A. D. 1998. *Strategies in learning and using a second language*. London: Longman.

Cohen, A. D. 2011. *Strategies in learning and using a second language*, 2nd edn. Harlow, UK: Longman Applied Linguistics/Pearson.

Ericsson, K. A., & H. A. Simon. 1984. *Protocol analysis: Verbal reports as data*. Cambridge, Mass.: MIT Press.

Ericsson, K. A., & H. A. Simon. 1993. *Protocol analysis: Verbal reports as data* Revised edn. Cambridge, MA: MIT Press.

Færch, C. & G. Kasper. 1987. From product to process – introspective methods in second language research. In C. Færch & G. Kasper (eds.), *Introspection in second language research*, 5–23. Clevedon: Multilingual Matters.

Field, J. 2008. *Listening in the language classroom*. Cambridge: Cambridge University Press.

Fox, M. C., K. A. Ericsson & R. Best. 2011. Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin* 137. 316–344. doi.org/10.1037/a0021663

Goh, C. C. M. 1998. How ESL learners with different listening abilities use comprehension strategies and tactics. *Language Teaching Research* 2. 124–147. doi.org/10.1177/136216889800200203

Goh, C. C. M. 2002. Exploring listening comprehension tactics and their interaction patterns. *System* 30. 185–206. doi.org/10.1016/s0346-251x(02)00004–0

Graham, S. 1997. *Effective language learning: Positive strategies for advanced level language learning*. Clevedon Avon, UK: Multilingual Matters Ltd.

Graham, S., D. Santos & R. Vanderplank. 2008. Listening comprehension and strategy use: A longitudinal exploration. *System* 36. 52–68. doi.org/10.1016/j.system.2007.11.001

Graham, S., D. Santos & R. Vanderplank. 2011. Exploring the relationship between listening development and strategy use. *Language Teaching Research* 15. 435–456. doi.org/10.1177/1362168811412026

Green, A. 1998. *Verbal protocol analysis in language testing research: A handbook*. Cambridge: Cambridge University Press.

Gruba, P. 1999. *The role of digital video media in second language listening comprehension*. Unpublished Doctoral Dissertation. The University of Melbourne, Australia.

Jourdenais, R. 2001. Cognition, instruction and protocol analysis. In P. Robinson (Ed.), *Cognition and second language instruction*. 354–375. Cambridge: Cambridge University Press.

Kasper, G. 1998. Analysing verbal protocols. *TESOL Quarterly* 32. 358–362. doi.org/10.2307/3587591

Kim, H. S. 2002. We talk, therefore we think? A cultural analysis of the effect of talking on thinking. *Journal of Personality and Social Psychology* 83. 828–842. doi.org/10.1037/0022–3514.83.4.828

Laviosa, F. 1991. *A preliminary investigation of the listening problem-solving processes and strategies of five advanced learners of Italian as a second language*. Unpublished Doctoral Dissertation. State University of New York at Buffalo, USA.

Lincoln, Y. S. & E. G. Guba. 1985. *Naturalistic inquiry*. Beverly Hills, CA: Sage Publications.

Macaro, E., S. Graham & R. Vanderplank. 2007. A review of listening strategies: focus on sources of knowledge and on success. In A. D. Cohen & E. Macaro (eds.), *Language learner strategies: 30 years of research and practice*, 165–185. Oxford: Oxford University Press.

Mareschal, C. 2007. *Student perceptions of a self-regulatory approach to second language listening comprehension development*. Unpublished Doctoral Dissertation. University of Ottawa, Canada.

Miles, M. B., & A. M. Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*, 2nd edn. Thousand Oaks: Sage Publications.

Murphy, J. M. 1985. *An investigation into the listening strategies of ESL college students*. ERIC Document Reproduction Service (No. ED 278 275).

O'Malley, J. M., A. U. Chamot & L. Küpper. 1989. Listening comprehension strategies in second language listening. *Applied Linguistics* 10. 418–437.

Patton, M. Q. 2015. *Qualitative research and evaluation methods: Integrating theory and practice*, 4th edn. Thousand Oaks, CA: Sage.

Pressley, M. & P. Afflerbach. 1995. *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Roever, C. & Y. C. Pan. 2008. Test reviews. GEPT: General english proficiency test. *Language Testing* 25. 403–418. doi.org/10.1177/0265532208090159

Rost, M. 2001. Listening. In R. Carter & D. Nunan (eds.), *The Cambridge guide to teaching English to speakers of other languages*, 7–13. Cambridge: Cambridge University Press.

Vandergrift, L. 1997. The strategies of second language (French) listeners. *Foreign Language Annals* 30. 387–409. doi.org/10.1111/j.1944-9720.1997.tb02362.x

Vandergrift, L. 1998a. Constructing meaning in L2 listening: Evidence from protocols. In S. Lapkin (ed.), *French second language education in Canada: Empirical studies*, 89–119. Toronto: University of Toronto Press.

Vandergrift, L. 1998b. Successful and less successful listeners in French: What are the strategy differences? *The French Review* 71(3). 370–395.

Vandergrift, L. 2003. Orchestrating strategy use: Toward a model of the skilled second language listener. *Language Learning*, 53, 463–496. doi.org/10.1111/1467–9922.00232

Wagner, E. 2008. Video listening tests: What are they measuring? *Language Assessment Quarterly* 5. 218–243.

Wu, Y. 1998. What do tests of listening comprehension test? – A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing* 15. 21–44. doi.org/10.1191/026553298673885021

Yeldham, M. 2009. *Approaches to second language listening instruction: Investigating the 'top-down/bottom-up' debate*. Unpublished Doctoral Dissertation. The University of Melbourne, Australia.

Yeldham, M., & R. T.-H. Chen. 2014. Conducting verbal reports to study Chinese-speaking English learners' listening: The use of prompts. *The European Journal of Applied Linguistics and TEFL* 3(1). 55–74.

Yeldham, M., & P. Gruba. 2014. Toward an instructional approach to developing interactive second language listening. *Language Teaching Research* 18. 33–53. doi:10.1177/1362168813505395

Yeldham, M., & P. Gruba. 2016. The development of individual learners in an L2 listening strategies course. *Language Teaching Research* 20. 9–34. doi:10.1177/1362168814541723

Young, M. Y. C. 1997. A serial ordering of listening comprehension strategies used by advanced ESL learners in Hong Kong. *Asian Journal of English Language Teaching* 7. 35–53.

Appendix A. VR task directions, and sample VR questions

(a)   VR warm-up task directives:

Soon we are going to get you to tell us your thoughts while you listen. But first, to get you used to verbalizing your thoughts, please tell us out loud how you work out the math problem, 32 multiplied by 12.

Now imagine you are walking through your house, starting from the front door, and tell us what you 'see' as you do this.

(b)   Pre-VR Directive:

Now we're going to play some listening passages to you. We will regularly stop each passage after every sentence or two, and each time we stop it, we'd like you to tell us *what you heard*, and also *what you were thinking while you were listening*. Please tell us everything that was in your mind; please give us as full and as accurate an account of your thoughts as you can.

    (Note: prior to using style 4, the section, *what you were thinking while you were listening*, was replaced with, *how you tried to understand while you were listening*.)

(c)   VR questions used to probe learner thoughts:

    VR style 1 (typical questions): Why did you say ___?; Why did you think ____?; How did you know ___?
    VR style 2: (Similar questions to style 1 were used, followed by): Was there anything else you were thinking that you forgot to tell me?
    VR style 3: (No questions were asked. However, when this style was used following styles 1 and 2, this directive was given immediately before use of this style 3: "With the next passage(s), we won't ask you any questions. Please just tell us what you heard, and what you were thinking while you were listening.")

Appendix B. Sample interview questions

Questions asked to investigate the learners' views of mediation styles 1–3
    (First we described each style to the learner, then asked him/her):

1.   Please compare these three styles: how did you feel about each one?
2.   Which style did you like the most, and why?
3.   Which style did you like the least, and why?

Questions asked to investigate the learners' comparison of the two pre-VR instruction styles:
1.   Did you notice a difference between the two styles? (If so,) What was the difference?
2.   Did you feel that the styles had different effects on your VR?